THE QUARTERLY JOURNAL OF MATHEMATICS

OXFORD SECOND SERIES

Volume 12 *No.* 48 *December* 1961

OXFORD
AT THE CLARENDON PRESS

Price 16*s. net*

PRINTED IN GREAT BRITAIN BY VIVIAN RIDLER AT THE UNIVERSITY PRESS, OXFORD

THE QUARTERLY JOURNAL OF MATHEMATICS

OXFORD SECOND SERIES

Edited by T. W. CHAUNDY, U. S. HASLAM-JONES, E. C. THOMPSON

THE QUARTERLY JOURNAL OF MATHEMATICS (OXFORD SECOND SERIES) is published at 16*s.* net for a single number with an annual subscription (for four numbers) of 55*s.* post free.

Papers, of a length normally not exceeding 20 printed pages of the Journal, are invited on subjects of Pure and Applied Mathematics, and should be addressed 'The Editors, Quarterly Journal of Mathematics, Clarendon Press, Oxford.' *Authors are referred to 'The Printing of Mathematics' (Oxford University Press, 1954) for detailed advice on the preparation of mathematical papers for publication.* The Editors as a rule will not wish to accept material that they cannot see their way to publish within a twelvemonth.

While every care is taken of manuscripts submitted for publication, the Publisher and the Editors cannot hold themselves responsible for any loss or damage. Authors are advised to retain a copy of anything they may send for publication. Authors of papers printed in the Quarterly Journal will be entitled to 50 free offprints.

Correspondence on the *subject-matter* of the Quarterly Journal should be addressed, as above, to 'The Editors', at the Clarendon Press. All other correspondence should be addressed to the publishers:

OXFORD UNIVERSITY PRESS

AMEN HOUSE, LONDON, E.C. 4

HIGHER WHITEHEAD PRODUCTS

By K. A. HARDIE (*Rondebosch, S.A.*)

[Received 18 March 1960.—Revised 25 March 1961]

1. Introduction

LET $\mathscr{X} = ({}_1X, {}_2X,..., {}_jX)$ be an ordered set of Hausdorff topological spaces. We denote all base-points by $*$ and assume that the base-point of any topological product is the point whose coordinates are all base-points. Let $\alpha_i \in \pi_{r_i}({}_iX, *)$ $(r_i \geqslant 1;\ i = 1, 2,..., j)$ be elements and let $r = \sum r_i$. If $j = 2$ and ${}_1X = {}_2X = X$, the classical Whitehead product $[\alpha_1, \alpha_2]$ is an element of $\pi_{r-1}(X)$. If $j \geqslant 2$, an *exterior Whitehead product* can readily be defined. Let $\prod \mathscr{X}$ be the topological product ${}_1X \times {}_2X \times ... \times {}_jX$ and let $\prod^{\cdot} \mathscr{X}$ be the subset whose points have at least one coordinate at a base-point. Then the exterior Whitehead product $\{\alpha_1, \alpha_2,..., \alpha_j\}$ we define below to be a certain element of $\pi_{r-1}(\prod^{\cdot} \mathscr{X}, *)$. There appears to be no entirely satisfactory definition of interior products if $j > 2$. Let $\mathscr{S}$ be the ordered set $(S^{r_1}, S^{r_2},..., S^{r_j})$ and let $\iota_n \in \pi_n(S^n)$ be the class of the identity map. Then, if

$${}_1X = {}_2X = ... = {}_jX = X,$$

one possibility is to define

$$[\alpha_1, \alpha_2,..., a_j]' \in \pi_{r-1}(X, *)$$

to be the set of elements $f_*\{\iota_{r_1}, \iota_{r_2},..., \iota_{r_j}\}$ for all possible maps

$$f\colon \textstyle\prod^{\cdot} \mathscr{S}, * \rightarrow X, *$$

whose restriction to the 'subset' S^{r_i} of $\prod^{\cdot} \mathscr{S}$ is a representative of α_i $(i = 1, 2,..., j)$. If $j = 3$, we obtain the construction due to E. C. Zeeman [see (2)]. However, if $j > 2$, $[\alpha_1, \alpha_2,..., \alpha_j]'$ may be empty, and, if $j > 3$, the modulus of the construction has yet to be determined.

An alternative procedure is to embed X in the subspace X_{j-1} of the reduced product space X_∞ of X [see (5)]. Then a preferred map $f\colon \prod^{\cdot} \mathscr{S} \rightarrow X_{j-1}$ can always be defined. Somewhat more generally consider the case ${}_iX = X_{m_i}$ $(i = 1, 2,..., j)$: let $m' = \sum m_i$, $m = m' - \min(m_i)$ and let $f\colon \prod^{\cdot} \mathscr{X} \rightarrow X_m$ be the map which agrees with the natural map $p\colon \prod \mathscr{X} \rightarrow X_{m'}$ [see (5) 176]. We describe

$$[\alpha_1, \alpha_2,..., \alpha_j] = f_*\{\alpha_1, \alpha_2,..., \alpha_j\} \in \pi_{r-1}(X_m, *) \tag{1.1}$$

as the *inner Whitehead product* of $\alpha_1, \alpha_2, \ldots, \alpha_j$. If $j = 2$ and $m_1 = m_2$, then we recover the classical Whitehead product.

Let $X = A$ be a CW-complex with but one 0-cell, namely $*$, and let sA be the (reduced) suspension of A. In (**4**) we considered a generalization of the Hopf construction which associates with a based map $F: \prod \mathscr{S} \to A_\infty$ an element $c(F) \in \pi_{r+1}(sA)$. Here $c(F)$ was shown to possess certain properties when F is a map strongly of type $(\alpha_1, \alpha_2, \ldots, \alpha_j)^{m'-1}$. We recall that this means (in effect) that $F(\prod \mathscr{S}) \subseteq A_{m'-1}$ and that F agrees on $\prod^{\cdot} \mathscr{S}$ with a map

$$F' = p(f_1 \times f_2 \times \ldots \times f_j): \prod \mathscr{S} \to A_{m'}, \tag{1.2}$$

where $f_i: S^{r_i}, * \to A_{m_i}, *$ represents $\alpha_i \in \pi_{r_i}(A_{m_i})$ $(i = 1, 2, \ldots, j)$. It is an easy consequence of the definitions that the injection of $[\alpha_1, \alpha_2, \ldots, \alpha_j]$ in $\pi_{r-1}(A_{m'-1})$ is the obstruction to the existence of a map strongly of type $(\alpha_1, \alpha_2, \ldots, \alpha_j)^{m'-1}$.

If $\alpha_1 = \alpha_2 = \ldots = \alpha_j = \alpha$, we denote $[\alpha_1, \alpha_2, \ldots, \alpha_j]$ by $[\alpha]^j$. Let $\eta_n \in \pi_{n+1}(S^n)$ $(n \geqslant 2)$ be generators. We shall prove the theorem:

THEOREM 1.3. (*a*) *If $n > 1$ and if $j \neq 2$ when n is odd, then $[\iota_n]^j$ is an element of infinite order.*

(*b*) $[\eta_2 \circ \eta_3 \circ \eta_4, \iota_2, \iota_2, \iota_2] = 0 \in \pi_{10}(S^2_3)$.

(*c*) *If p is an odd prime and $\beta_n = [\iota_n, [\iota_n]^{p-1}] \in \pi_{pn-2}(S_{p-2})$, then*

(i) $\beta_2 = 0$ *and* (ii) *for $m > 1$, β_{2m} is of order p.*

(1.3)(*b*) implies the existence of a map F strongly of type

$$(\eta_2 \circ \eta_3 \circ \eta_4, \iota_2, \iota_2, \iota_2)^3.$$

From 1.9 and 1.10 of (**4**) we obtain immediately the corollary:

COROLLARY 1.4. *$c(F)$ is a non-zero element of $\pi_{12}(S^3)$. The (James) reduced product filtration of $c(F)$ is 4.*

Now let F be a map strongly of type $(\iota_2, [\iota_2]^{p-1})^{p-2}$. We shall prove the theorem:

THEOREM 1.5. *$c(F)$ is an element of order p in $\pi_{2p}(S^3)$.*

In § 2 I mention a number of elementary properties of the higher products including a result which in a special case reduces to the familiar Jacobi identity.

2. Elementary properties

Let V^n and S^n denote the n-element and n-sphere as represented in (**6**). We shall assume that orientations for V^n, S^n and products of

spheres have been chosen as in (6). Let

$$\psi\colon V^r, S^{r-1}, * \to \prod \mathscr{S}, \prod^{\cdot} \mathscr{S}, *$$

be an orientation-preserving characteristic map for the r-cell of $\prod \mathscr{S}$ and let

$$\psi|\colon S^{r-1}, * \to \prod^{\cdot} \mathscr{S}, *$$

be the map which agrees with ψ. Let $g_i\colon S^{r_i}, * \to {}_iX, *$ be a representative of $\alpha_i \in \pi_{r_i}({}_iX)$ $(i = 1, 2, \ldots, j)$, and let $g\colon \prod^{\cdot} \mathscr{S} \to \prod^{\cdot} \mathscr{X}$ be the map which agrees with the product map $g_1 \times g_2 \times \ldots \times g_j\colon \prod \mathscr{S} \to \prod \mathscr{X}$. We define

$$\{\alpha_1, \alpha_2, \ldots, \alpha_j\} = g_*\{\psi|\} \in \pi_{r-1}(\textstyle\prod^{\cdot} \mathscr{X}, *). \tag{2.1}$$

The exterior Whitehead product possesses an obvious naturality property. We also observe that, if

$$d\colon \pi_r(\textstyle\prod \mathscr{X}, \prod^{\cdot} \mathscr{X}, *) \to \pi_{r-1}(\prod^{\cdot} \mathscr{X}, *)$$

is the homotopy boundary homomorphism and if $\star$ is the product defined in § 5 of (1), then we have the theorem:

THEOREM 2.2. *If $r_i \geqslant 2$ $(i = 1, 2, \ldots, j)$, then*

$$\{\alpha_1, \alpha_2, \ldots, \alpha_j\} = d((\ldots((\alpha_1 \star \alpha_2) \star \alpha_3) \star \ldots) \star \alpha_j).$$

Commutation rules and j-linearity (in the case $r_i \geqslant 2$ $(i = 1, 2, \ldots, j)$) follow from the corresponding properties of the $\star$ product. Somewhat less obvious is the following relation for $j \geqslant 3$. Let $\mathscr{X}_i$ be the ordered set $\mathscr{X} - ({}_iX)$, let $\prod^{\cdot\cdot} \mathscr{X} \subseteq \prod \mathscr{X}$ be the subset whose points have at least two coordinates at base-points and let

$$\theta_i\colon \pi_n({}_iX) \to \pi_n(\textstyle\prod^{\cdot\cdot} \mathscr{X}), \qquad \phi_i\colon \pi_n(\prod^{\cdot} \mathscr{X}_i) \to \pi_n(\prod^{\cdot\cdot} \mathscr{X})$$

be the homomorphisms induced by the canonical injections. We have the theorem:

THEOREM 2.3. *Let $\eta(i) = r_i(r_1 + r_2 + \ldots + r_i) + 1$ $(r_i \geqslant 2;\ i = 1, 2, \ldots, j)$, then*

$$\sum_{i=1}^{j} (-1)^{\eta(i)} [\theta_i \alpha_i, \phi_i\{\alpha_1, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_j\}] = 0 \in \pi_{r-2}(\textstyle\prod^{\cdot\cdot} \mathscr{X}).$$

By applying the boundary operator to the formula of Nakaoka and Toda [(9) 5] and using 3.5 of (1) we obtain the result for the case of the universal example.

Let E denote the suspension homomorphism and let

$$\delta_i \in \pi_{q_i - 1}(S^{r_i - 1}) \quad (i = 1, 2, \ldots, j).$$

The proof of G. W. Whitehead's identity 3.59 of (**13**) can obviously be extended to yield

$$\{\alpha_1 \circ E\delta_1, \alpha_2 \circ E\delta_2,..., \alpha_j \circ E\delta_j\} = \{\alpha_1, \alpha_2,..., \alpha_j\} \circ \delta_1 * \delta_2 * ... * \delta_j. \tag{2.4}$$

The corresponding relations satisfied by the inner products are easily obtained. For example we have the commutation rule

$$[\alpha_1, \alpha_2,..., \alpha_{i-1}, \alpha_{i+1}, \alpha_i, \alpha_{i+2},..., \alpha_j] = (-1)^{r_i r_{i+1}}[\alpha_1, \alpha_2,..., \alpha_j], \tag{2.5}$$

and, if $r_i \geqslant 2$ $(i = 1, 2,..., j)$, the identity

$$\sum_{i=1}^{j} (-1)^{\eta(i)}\phi_i[\alpha_i, [\alpha_1,..., \alpha_{i-1}, \alpha_{i+1},..., \alpha_j]] = 0 \in \pi_{r-2}(X_{m''}), \tag{2.6}$$

where $m'' = \sum m_i - \min(m_i + m_k)$ $(i \neq k)$ and where ϕ_i denotes the appropriate injection homomorphism. Let $h\colon X \to Y$ be a based map and let $h_\infty\colon X_\infty \to Y_\infty$ map termwise by h [see (**5**) 172]. Then, if for each $n \geqslant 0$ we denote by $h_*\colon \pi_q(X_n) \to \pi_q(Y_n)$ the homomorphisms induced by maps which agree with h_∞, we have

$$h_*[\alpha_1, \alpha_2,..., \alpha_j] = [h_* \alpha_1, h_* \alpha_2,..., h_* \alpha_j]. \tag{2.7}$$

Finally we remark that (2.4) yields

$$[\alpha_1 \circ E\delta_1, \alpha_2 \circ E\delta_2,..., \alpha_j \circ E\delta_j] = [\alpha_1, \alpha_2,..., \alpha_j] \circ \delta_1 * \delta_2 * ... * \delta_j. \tag{2.8}$$

3. Computations

We require some additional notation. Let r be an integer and let

$${}_rh\colon S^n_\infty, S^n_{r-1} \to S^{rn}_\infty, *$$

be the combinatorial extension of a map $S^n_r \to S^{rn}$ which shrinks S^n_{r-1} and is of degree one on the rn-cell [see (**5**)]. Let

$${}_rh'\colon S^n_{2r-1}, S^n_{r-1} \to S^{rn}, *$$

be the map which agrees with ${}_rh$. We shall also denote by ${}_rh$, ${}_rh'$ the homomorphisms

$$\pi_i(S^n_\infty, S^n_{r-1}) \to \pi_i(S^{rn}_\infty), \qquad \pi_i(S^n_{2r-1}, S^n_{r-1}) \to \pi_i(S^{rn})$$

which they induce. Let $\phi\colon \pi_i(S^n_\infty) \to \pi_{i+1}(S^{n+1})$ be the canonical isomorphism of (**5**). We recall that the Hopf–James invariants

$$H_r\colon \pi_i(S^{n+1}) \to \pi_i(S^{rn+1})$$

are the homomorphisms $H_r = \phi\, {}_rh\phi^{-1}$.

Let X be a space. Then sX the (reduced) *suspension* of X is the space obtained from $X \times I$ by shrinking the subset

$$X \times ((0) \cup (1)) \cup (*) \times I$$

to $*$. If $f\colon X \to Y$ is a based map, let $s(f)\colon sX \to sY$ be the map such that

$$s(f)\nu(x,t) = \nu(fx,t),$$

where $x \in X$, $t \in I$, and ν denotes the suspension identification. Let Ω be the space of loops on S^{n+1} and let $\phi_1\colon S^n_\infty \to \Omega$ be a canonical map [see (**5**)]. Let

$$\psi\colon sS^n_\infty \to S^{n+1} \tag{3.1}$$

be the map such that $\psi\nu(x,t) = \phi_1 x(t)$, where $x \in S^n_\infty$, $t \in I$.

Proof of Theorem 1.3 (*a*). Suppose that $n > 1$, $j \neq 2$ if n is odd, and that $k[\iota_n]^j = 0$. Then the j-linearity implies the existence of a map F strongly of type $(k\iota_n, \iota_n, \ldots, \iota_n)^{j-1}$ and $c(F) \in \pi_{jn+1}(S^{n+1})$. By (1.10) of (**4**), $H_j c(F) = k\iota_{jn+1}$. Thus $c(F)$ is an element of infinite order, contradicting the result given by J.-P. Serre. If $n = 1$, the argument fails, for we no longer have j-linearity.

Proof of Theorem 1.3 (*b*). Let

$$\beta = [\eta_2 \circ \eta_3 \circ \eta_4, \iota_2, \iota_2, \iota_2] \in \pi_{10}(S^2_3).$$

Applying (2.8) we have

$$\beta = [\eta_2, \iota_2, \iota_2, \iota_2] \circ \eta_8 \circ \eta_9,$$

so that $2\beta = 0$. Since $\pi_{10}(S^2) = Z_{15}$, the exactness of the sequence

$$\to \pi_{10}(S^2) \xrightarrow{i} \pi_{10}(S^2_3) \xrightarrow{j} \pi_{10}(S^2_3, S^2) \xrightarrow{d} \pi_9(S^2) \to$$

implies that it will be sufficient to prove that $j\beta = 0$. By (1.2) of (**3**),

$${}_2h'\colon \pi_i(S^2_3, S^2) \to \pi_i(S^4)$$

is a $\mathscr{C}_2$-isomorphism, so that it will be sufficient to prove that ${}_2h'j\beta = 0$. We have

$${}_2h'j[\eta_2, \iota_2, \iota_2, \iota_2] \in \pi_8(S^4) \approx Z_2 + Z_2,$$

generated by $\eta_4 \circ \nu_5$ and $\nu_4 \circ \eta_7$ [see (**10**)]. Here ν_4 denotes the Hopf element and ν_5, ν_6, etc., its successive suspensions. Now $E(\nu_4 \circ \eta_7) \neq 0$ and thus

$${}_2h'j[\eta_2, \iota_2, \iota_2, \iota_2] = ({}_2h')_*[\eta_2, \iota_2, \iota_2, \iota_2] \neq \nu_4 \circ \eta_7,$$

for, if in the diagram

$$\begin{array}{ccc} S^2_3 & \xrightarrow{k} & S^2_\infty \\ \downarrow {}_2h' & & \downarrow {}_2h \\ S^4 & \xrightarrow{k'} & S^4_\infty \end{array}$$

k and k' are injection maps, then $[\eta_2, \iota_2, \iota_2, \iota_2]$ belongs to the kernel of k_* while $(k')_*$ is equivalent to E. If

$${}_2h'j[\eta_2, \iota_2, \iota_2, \iota_2] = 0,$$

then 1.3 (*b*) is proved. Suppose that

$$ {}_2h'j[\eta_2, \iota_2, \iota_2, \iota_2] = \eta_4 \circ \nu_5. $$

Then $$ {}_2h'j\beta = E(\eta_3 \circ \nu_4 \circ \eta_7 \circ \eta_8) = 0 $$

since $$ \eta_3 \circ \nu_4 \circ \eta_7 \circ \eta_8 \in \pi_9(S^3) \approx Z_3. $$

Proof of Theorem 1.3 (*c*) (i). Let

$$ g'\colon V^{2p-2}, S^{2p-3} \to S^2_{p-1}, S^2_{p-2} $$

be a characteristic map for the $2(p-1)$-cell of S^2_{p-1} and let

$$ g\colon S^{2p-3} \to S^2_{p-2} $$

agree with g'. Then $\{g\} = [\iota_2]^{p-1}$ and, if $\mu\colon S^2 \to S^2_\infty$ is the injection and $[\mu_* \iota_2, \{g'\}]''$ is the Blakers–Massey Whitehead product, we have

$$ d[\mu_* \iota_2, \{g'\}]'' = \pm\beta_2 $$

by (3.5) of (**1**), where d is the boundary homomorphism in the exact sequence

$$ \to \pi_r(S^2_{p-2}) \xrightarrow{i} \pi_r(S^2_{p-1}) \xrightarrow{j} \pi_r(S^2_{p-1}, S^2_{p-2}) \xrightarrow{d} \pi_{r-1}(S^2_{p-2}) \to. $$

It will be sufficient to prove that

$$ [\mu_* \iota_2, \{g'\}]'' \in j\pi_{2p-1}(S^2_{p-1}). $$

By (5.8)″ of (**11**) we have a direct sum decomposition

$$ \pi_{2p-1}(S^2_{p-1}, S^2_{p-2}) \approx (g')_* \pi_{2p-1}(V^{2p-2}, S^{2p-3}) + [\pi_2(S^2_{p-1}), \{g'\}]''. \quad (3.2) $$

Let $\alpha \in \pi_{2p}(S^3)$ be an element of order p. By (1.3) of (**7**) there exists an element $\gamma \in \pi_{2p-1}(S^2_{p-1})$ such that $j\gamma \neq 0$ and such that $\alpha = \phi k_* \gamma$, where $k\colon S^2_{p-1} \to S^2_\infty$ is the injection. Then the order of γ is either infinite or a multiple of p. Since

$$ \pi_{2p-1}(V^{2p-2}, S^{2p-3}) \approx \pi_{2p-2}(S^{2p-3}) \approx Z_2 $$

and $[\pi_2(S^2_{p-1}), \{g'\}]''$ is cyclic, (3.2) implies that

$$ 2j\gamma = m[\mu_* \iota_2, \{g'\}]'' $$

for some integer m. Furthermore, by (5.9)″ of (**11**),

$$ j[\iota_2]^p = p[\mu_* \iota_2, \{g'\}]''. $$

It follows that m is prime to p, for, if $m = qp$, then

$$ j(2\gamma - q[\iota_2]^p) = 0, $$

which implies that $$ 2\gamma - q[\iota_2]^p \in i\pi_r(S^2_{p-2}), $$

and hence that $$ \phi k_*(2\gamma - q[\iota_2]^p) = 2\phi k_* \gamma = 2\alpha $$

has filtration $p-2$ contradicting (1.3) of (7). We may deduce that

$$[\mu_* \iota_2, \{g'\}]'' \in j\pi_{2p-1}(S^2_{p-1}),$$

which completes the proof.

Proof of Theorem 1.3 (*c*) (ii). The case $p = 3$ of the following argument is due to Nakaoka and Toda (9). Let n be even. By (2.6) with $j = p$, $\alpha_1 = \alpha_2 = \ldots = \alpha_p = \iota_n$, or by applying the boundary operator to (5.9)″ of (11), we have $p\beta_n = 0$. It remains to prove that $\beta_{2m} \neq 0$ if $m \geqslant 2$. Let

$$g'\colon V^{n(p-1)}, S^{n(p-1)-1} \to S^n_{p-1}, S^n_{p-2}$$

be a characteristic map for the $n(p-1)$-cell of S^n_{p-1} and let

$$g\colon S^{n(p-1)-1} \to S^n_{p-2}$$

agree with g'. Suppose that $\beta_n = 0$. Then there exists a map

$$F\colon S^n \times S^{n(p-1)-1} \to S^n_\infty$$

strongly of type $(\iota_n, [\iota_n]^{p-1})^{p-2}$ such that $F(*, x) = gx$ if $x \in S^{n(p-1)-1}$. It follows that there exists a map

$$f'\colon S^n \times S^{n(p-1)-1} \cup (*) \times V^{n(p-1)} \to S^n_{p-1}$$

which agrees with F and is such that $f'(*, x) = g'x$ if $x \in V^{n(p-1)}$. Let $K = S^n_{p-1} \cup \mathscr{E}^{pn}$, where $\mathscr{E}^{pn}$ is the pn-cell of $S^n \times V^{n(p-1)}$ attached by the map

$$f\colon S^n \times V^{n(p-1)}, S^n \times S^{n(p-1)-1} \cup (*) \times V^{n(p-1)}, S^n \times S^{n(p-1)-1}$$
$$\to K, S^n_{p-1}, S^n_{p-1},$$

which agrees with f'. Let $e_i \in H^{in}(K)$ $(i = 1, 2, \ldots, p)$ be generators such that the j-fold cup product

$$(e_1)^j = j!\, e_j \quad (1 \leqslant j \leqslant p-1)$$

[see (2.4) of (11)]. Let d_{p-1}, d_p be generators of

$$H^{n(p-1)}(K, S^n_{p-2}),\ H^{pn}(K, S^n_{p-2})$$

respectively and let f^* denote the various cohomology and relative cohomology homomorphisms induced by f. Then f^*e_1, f^*d_{p-1}, f^*d_p generate

$$H^n(S^n \times V^{n(p-1)}), \quad H^{n(p-1)}(S^n \times V^{n(p-1)}, S^n \times S^{n(p-1)-1}),$$
$$H^{pn}(S^n \times V^{n(p-1)}, S^n \times S^{n(p-1)-1})$$

respectively, so that

$$f^*e_1 . f^*d_{p-1} = \pm f^*d_p,$$

the full point denoting the relative cup product. Since

$$f^*\colon H^{pn}(K, S^n_{p-2}) \to H^{pn}(S^n \times V^{n(p-1)}, S^n \times S^{n(p-1)-1})$$

is an isomorphism, we have $e_1 . d_{p-1} = \pm d_p$ and hence $e_1 . e_{p-1} = \pm e_p$. Therefore

$$(e_1)^p = e_1 . (e_1)^{p-1} = (p-1)!\, e_1 . e_{p-1} = \pm (p-1)!\, e_p \equiv \pm e_p \pmod p$$

by Wilson's theorem. It follows that the Steenrod pth power

$$\mathscr{P}^{\frac{1}{2}n} e_1 = (e_1)^p \neq 0.$$

Let $Y = S^{n+1} \cup \mathscr{E}^{pn+1}$, where $\mathscr{E}^{pn+1}$ is the $(pn+1)$ cell of sK attached by the map

$$\psi' \colon sK,\ s(S^n_{p-1}) \to Y,\ S^{n+1}$$

which agrees on $s(S^n_{p-1})$ with the map ψ of (3.1). By the naturality of the reduced power operations and the fact that they commute with suspension we have an isomorphism

$$\mathscr{P}^{\frac{1}{2}n} \colon H^{n+1}(Y, Z_p) \to H^{pn+1}(Y, Z_p).$$

Hence by Lemma 2.1 of (**12**) the mod p Hopf invariant homomorphism is non-trivial [see (**12**) 144]. Since by Theorem 5 of (**8**) this is only the case if $n = 2$, we have the required contradiction.

Proof of Theorem 1.5. Let $n = 2$. Then the construction in the proof of (1.3) (*c*) (ii) is possible and the $(2p+1)$-cell of $Y = S^3 \cup \mathscr{E}^{2p+1}$ is attached by a map whose class $\alpha \in \pi_{2p}(S^3)$ is of order p. We shall prove that $\alpha = \pm c(F)$. Let Q be the space

$$S^2 \times S^{2p-3} \cup (*) \times V^{2p-2}$$

and let $\gamma \in \pi_{2p-1}(Q)$ be the class of the attaching map of the $2p$-cell of $S^2 \times V^{2p-2}$. Then

$$\alpha = (\psi')_* E(f')_* \gamma = (\psi')_* (s(f'))_* E\gamma. \tag{3.3}$$

Consider the commutative diagram

$$\begin{tikzcd}
W = S^3 \vee S^{2p-2} \vee S^{2p} \arrow[d, "k''"] & s(S^2 \times S^{2p-3}) \arrow[l, "\theta"'] \arrow[r, "s(F)"] \arrow[d, "k"] & sS^2_\infty \arrow[r, "\psi"] & S^3 \\
R = S^3 \vee V^{2p-1} \vee S^{2p} & sQ \arrow[l, "\theta'"'] \arrow[r, "s(f')"] & sS^2_{p-1}, \arrow[u, "k'"'] \arrow[ur, "\psi'"'] &
\end{tikzcd} \tag{3.4}$$

where θ is the homotopy equivalence of (1.5) of (**4**), where θ' agrees with θ and maps the interior of the $(2p-1)$ cell homeomorphically, and where k, k', and k'' are injections. Let $\iota \in \pi_{2p}(W)$ be the class of the injection map $S^{2p} \to W$. Then we recall [(**4**) (1.6)] that

$$c(F) = \psi_*(s(F))_*(\theta_*)^{-1}\iota. \tag{3.5}$$

Since θ' is also a homotopy equivalence, in view of (3.3), (3.5), and the commutativity of the corresponding diagram of induced homomor-

phisms it will be sufficient to prove that

$$(k'')_* \iota = \pm(\theta')_* E\gamma.$$

We have $$\pi_{2p}(R) \approx \pi_{2p}(S^3)+\pi_{2p}(S^{2p}),$$

the direct summands being embedded by injections; $(k'')_* \iota$ generates the summand $\pi_{2p}(S^3)$. Let δ_1 be the component of $(\theta')_* E\gamma$ in $\pi_{2p}(S^3)$. Then, if $\pi\colon Q \to S^2$ is the projection, $\delta_1 = (s(\pi))_* E\gamma$ by the definition of θ'. But we have $\pi_* \gamma = 0$ since π may be extended to the projection $S^2 \times V^{2p-2} \to S^2$, and therefore $\delta_1 = 0$. Let δ_2 be the component of $(\theta')_* E\gamma$ in $\pi_{2p}(S^{2p})$. Then $\delta_2 = \lambda \iota_{pn}$ for some integer $\lambda \neq 0$, for otherwise $s(S^2 \times V^{2p-2})$ would have the homotopy type of $S^3 \vee S^{2p} \vee S^{2p+1}$, which it has not. Further we assert that $\lambda = \pm 1$, for otherwise the homology groups of $s(S^2 \times V^{2p-2})$ would have torsion elements. This completes the proof of Theorem 1.5.

REFERENCES

1. A. L. Blakers and W. S. Massey, 'Products in homotopy theory', *Ann. of Math.* 58 (1953) 295–324.
2. K. A. Hardie, 'On a construction of E. C. Zeeman', *J. London Math. Soc.* 35 (1960) 452–64.
3. —— 'On the kernel of suspension', ibid., to appear.
4. —— 'A generalization of the Hopf construction', *Quart. J. of Math.* (Oxford), *supra* 196–204.
5. I. M. James, 'Reduced product spaces', *Ann. of Math.* 62 (1955) 170–97.
6. —— 'On the suspension triad', ibid. 63 (1956) 191–247.
7. —— 'Filtration of the homotopy groups of spheres', *Quart. J. of Math.* (Oxford) (2) 9 (1958) 301–9.
8. A. Liulevicius, 'The factorization of cyclic reduced powers', *Proc. Nat. Acad. Sci.* (*U.S.A.*) 46 (1960) 978–81.
9. M. Nakaoka and H. Toda, 'On Jacobi identity for Whitehead products', *J. Inst. Polytech. Osaka City Univ.* A 5 (1954) 1–13.
10. H. Toda, 'Generalised Whitehead products and homotopy groups of spheres', ibid. 3 (1952) 43–82.
11. —— 'On the double suspension', ibid. 7 (1956) 103–45.
12. —— 'p-primary components of homotopy groups II', *Memoirs of the College of Science, Univ. of Kyoto* A 31 (1958) 143–60.
13. G. W. Whitehead, 'A generalization of the Hopf invariant', *Ann. of Math.* 51 (1950) 192–237.

ON THE RATE OF CONVERGENCE TO THE CONNECTIVE CONSTANT OF THE HYPERCUBICAL LATTICE

By J. M. HAMMERSLEY (*Oxford*)

[Received 15 December 1960]

Statement and discussion of results

THE *hypercubical lattice* consists of all points with (positive, negative, or zero) integer coordinates in d-dimensional euclidean space, where d is fixed and $d \geqslant 2$. Throughout this paper, the term *point* will mean a point of the hypercubical lattice; and we shall write $(a, b,..., z)$ with or without suffixes for the coordinates of such a point. An *n-stepped self-avoiding walk* is an ordered sequence of $n+1$ mutually distinct points such that any two successive points of the sequence are unit distance apart. We write $S_n(A)$ for an n-stepped self-avoiding walk with A as its first point. Because of the homogeneity of the hypercubical lattice, the number of distinct $S_n(A)$ with a fixed A is independent of A, and will be denoted by $f(n)$. The function $\kappa(n)$ is defined by

$$f(n) = e^{n\kappa(n)}. \tag{1}$$

It is a particular case of results proved elsewhere (1) that

$$\kappa(n) \to \kappa \quad \text{as} \quad n \to \infty, \tag{2}$$

where the constant κ depends only on d and is called the *connective constant* of the hypercubical lattice. The function $f(n)$ appears in various branches of applied mathematics, for instance in problems connected with long-chain polymers, with percolation, with ferromagnetism, and with impurities in crystalline solids; and these applications have been hampered by lack of pure-mathematical knowledge of the properties of $f(n)$. The case $d = 3$ is the most important one.

In this paper I shall improve (2) by showing that

$$0 \leqslant \kappa(n)-\kappa \leqslant \frac{2d}{(d-1)n^{1/d}}\log\{dn(n+1)\}. \tag{3}$$

This result is presumably capable of considerable further improvement though I do not know how to effect this. Probably $\kappa(n)-\kappa$ behaves like a multiple of $n^{-1}\log n$ as $n \to \infty$. This behaviour is a simple

consequence of the conjecture that the ratio $f(n)/f(n-1)$ can be represented as one or other of two asymptotic series in $1/n$ according as n is even or odd. The empirical support for this conjecture is impressively good (**6**). Alternatively the same behaviour for $\kappa(n)-\kappa$ follows from the assumption (supported by certain heuristic arguments) that a similar pair of asymptotic series will represent the probability of initial ring closure (i.e. that a randomly chosen $S_n(A)$ passes through more than one nearest neighbour of A). The value of κ is not known exactly; but the best estimates (**4**, **6**) are that $e^{\kappa} = 2{\cdot}6390 \pm 0{\cdot}0005$ when $d = 2$, while $e^{\kappa} = 4{\cdot}152 \pm 0{\cdot}003$ when $d = 3$. The corresponding values of $\kappa(18)-\kappa = 0{\cdot}0652 \pm 0{\cdot}0002$ when $d = 2$, and of

$$\kappa(11)-\kappa = 0{\cdot}0811 \pm 0{\cdot}0007$$

when $d = 3$ are much less than the respective right-hand sides of (3), namely 6·1546 and 8·3383.

The function $f(n)$ has been tabulated by Sykes (**3**) for $n \leqslant 18$ when $d = 2$ and for $n \leqslant 11$ when $d = 3$. His results, which may suggest improvements in (3), are as follows:

n	$f(n)$ $(d = 2)$	$f(n)$ $(d = 3)$
1	4	6
2	12	30
3	36	150
4	100	726
5	284	3534
6	780	16926
7	2172	81390
8	5916	387966
9	16268	1853886
10	44100	8809878
11	120292	41933286
12	324932	
13	881500	
14	2374444	
15	6416596	
16	17245332	
17	46466676	
18	124658732	

Sykes's method is to write down a recurrence relation between the numbers of self-avoiding walks and the numbers of two other kinds of linear graphs, known on account of their shape as *tadpoles* and *dumbbells*. For small values of n it is easier to count these latter graphs than to count the self-avoiding walks. But, for large values of n, his relation is of no assistance. Work now in progress with an electronic computer at the National Physical Laboratory in England is likely to extend the

foregoing table somewhat though it will not reach large values of n. O'Flaherty (7) has made rather similar calculations in America. About four years ago rumours reached England that an anonymous Russian mathematician had found a recurrence relation by which $f(n)$ could be computed for arbitrarily large n; but nothing further has been heard of this.

Thus, as the position now stands, there is (in contrast to the conventional type of enumerative problem on linear graphs) no known exact expression for $f(n)$ from which an asymptotic approximation could be deduced. The only method so far available for handling $f(n)$ has been the theory of subadditive functions, a method which starts from inequalities rather than equations. A function $\phi(n)$ is said to be *subadditive* if it satisfies

$$\phi(m+n) \leqslant \phi(m)+\phi(n); \tag{4}$$

and, according to the fundamental theorem [(2) Theorem 6.6.1] on such functions, (4) implies

$$-\infty \leqslant \inf_{n>0} n^{-1}\phi(n) = \lim_{n\to\infty} n^{-1}\phi(n) < +\infty. \tag{5}$$

When the central terms of (5) are finite and denoted by ψ, (5) implies

$$\psi n \leqslant \phi(n) = \psi n+o(n) \quad \text{as} \quad n \to \infty. \tag{6}$$

As we shall see presently, (6) provides an easy proof of (2), but in itself is unable to give a sharper result than (2) because it is possible to find subadditive functions for which the rate of convergence of $\phi(n)/n$ to ψ is arbitrarily slow. For this reason there has seemed to be little hope of improving (2). The main interest of the present paper is accordingly less in the result (3) since it is presumably far from best-possible, and more in the technique which has now been found of surmounting this apparent impasse. In its essentials the technique is to combine the use of subadditive functions with superadditive functions, namely functions satisfying

$$\phi^*(m+n) \geqslant \phi^*(m)+\phi^*(n), \tag{7}$$

which implies

$$\psi^* n \geqslant \phi^*(n) = \psi^* n+o(n) \quad \text{as } n \to \infty. \tag{8}$$

In fact, (8) is merely the result of writing $\phi^* = -\phi$ and invoking (5). Now, if a function is both subadditive and superadditive, it must be a multiple of n. We cannot expect quite as much as this in the present problem because $\log f(n)$ is not a multiple of n. What we shall show, however, is that, while $\log f(n)$ is subadditive, it contains, as it were, a superadditive component; and, by playing off this superadditive

component against the subadditive whole, we can deduce that $\log f(n)$ differs from a linear function of n by an amount $O(n^{(d-1)/d}\log n)$.

The great merit of subadditive functions in applied mathematics is that they can be used in problems which are so difficult that one can only write down inequalities instead of equations. The technique described in outline above offers the prospect of sharpening the analysis of a variety of such intractable problems. A somewhat similar use of the same idea occurs in the recent proof (5) of a well-known conjecture on the number of lattice polygons.

Proof of results

The number of distinct $(m+n)$-stepped walks, which start at a fixed point A and which are self-avoiding for their first m steps and for their last n steps (although the last n steps may intersect the first m steps), is $f(m)f(n)$. These walks include all $S_{m+n}(A)$. Hence

$$f(m+n) \leqslant f(m)f(n), \tag{9}$$

which implies that $\log f(n)$ is subadditive. Since also

$$1 \leqslant f(n) \leqslant (2d)^n, \tag{10}$$

because each step can be chosen in at most $2d$ ways, we see that $n^{-1}\log f(n)$ is bounded; and (2) and the left-hand side of (3) follow at once from (5) and (6).

Let $A_i = (a_i, b_i,..., z_i)$, where $i = 0, 1,..., n$, be the points of an $S_n(A_0)$. If $a_0 < a_i \leqslant a_n$ for $i = 1, 2,..., n$, we shall call the $S_n(A_0)$ a *special* n-stepped self-avoiding walk and denote it by $S_n^*(A_0)$. Let $f^*(n)$ denote the number of distinct $S_n^*(A_0)$ for fixed A_0. Clearly

$$1 \leqslant f^*(n) \leqslant f(n). \tag{11}$$

If we add an $S_m^*(A_n)$ to the end of an $S_n^*(A_0)$, the result is an $S_{m+n}^*(A_0)$, so that

$$f^*(m+n) \geqslant f^*(m)f^*(n). \tag{12}$$

Thus $\log f^*(n)$ is a superadditive function of n, and $n^{-1}\log f^*(n)$ is bounded; and by (8) there exists a finite constant λ such that

$$e^{\lambda n} \geqslant f^*(n) = e^{\lambda n + o(n)} \quad \text{as} \quad n \to \infty. \tag{13}$$

By (1), (2), (11), and (13), we obtain $\lambda \leqslant \kappa$, so that

$$f^*(n) \leqslant e^{\kappa n}. \tag{14}$$

Let A_0 be a fixed point, and let $A_i = (a_i, b_i,..., z_i)$ $(i = 0, 1,..., n)$ be the points of an $S_n(A_0)$. Let

$$\alpha = \max_{0 \leqslant i \leqslant n} a_i, \qquad \alpha' = \min_{0 \leqslant i \leqslant n} a_i, \tag{15}$$

and define β, β',..., ζ, ζ' similarly in terms of b_i,..., z_i. Since all the points A_i are distinct, we must have

$$(\alpha-\alpha'+1)(\beta-\beta'+1)...(\zeta-\zeta'+1) \geqslant n+1. \tag{16}$$

Let C denote the class of $S_n(A_0)$ which are such that firstly

$$\alpha-\alpha'+1 > n^{1/d} \tag{17}$$

and secondly they visit the hyperplane $a = \alpha'$ before they visit the hyperplane $a = \alpha$. By symmetry and by (16), there are at least $f(n)/2d$ distinct members in C. We now confine our attention to members of C, and write A_ρ for the last point of $S_n(A_0)$ belonging to the hyperplane $a = \alpha$. Write A_σ for the last of those points of $S_n(A_0)$ on the hyperplane $a = \alpha'$ which is such that $\sigma < \rho$. Here, of course, the values of ρ and σ depend upon the particular member of C under consideration. We have

$$\rho-\sigma \geqslant \alpha-\alpha' > n^{1/d}-1, \tag{18}$$

and thus, because ρ and σ are integers,

$$\rho-\sigma \geqslant [n^{1/d}], \tag{19}$$

where $[n^{1/d}]$ is the integer part of $n^{1/d}$. Since A_ρ and A_σ are distinct members of A_0, A_1,..., A_n and $\sigma < \rho$, the number of distinct pairs of values available for ρ and σ is at most $\frac{1}{2}n(n+1)$. Hence we can find a subclass C^* of C with a *fixed* pair of values ρ and σ such that C^* contains at least $f(n)/dn(n+1)$ distinct members. We now confine attention to members of C^*.

From the definition of A_ρ and A_σ, it follows that the portion of $S_n(A_0)$ from A_σ to A_ρ is an $S^*_{\rho-\sigma}(A_\sigma)$; and therefore, for prescribed A_σ, this portion can be constructed in at most

$$f^*(\rho-\sigma) \leqslant e^{\kappa(\rho-\sigma)} \tag{20}$$

ways. Again, by the definition of A_ρ and A_σ, if the last part of $S_n(A_0)$ from A_ρ to A_n is translated bodily and added to the end of the first part of $S_n(A_0)$ from A_0 to A_σ, the result will be an $S_{n-\rho+\sigma}(A_0)$. The number of ways in which these two parts can be constructed is therefore at most

$$f(n-\rho+\sigma), \tag{21}$$

and, if the two parts are prescribed, then A_σ is consequently prescribed. Hence the number of members of C^* is at most

$$e^{\kappa(\rho-\sigma)}f(n-\rho+\sigma), \tag{22}$$

by virtue of (20) and (21). To be precise, there is the possibility that $n = \rho-\sigma$, so that the first and last parts of $S_n(A_0)$ referred to above

are empty; but this possibility can be covered by the convention that $f(0) = 1$. We deduce from (22) that

$$f(n)/dn(n+1) \leqslant e^{\kappa(\rho-\sigma)}f(n-\rho+\sigma). \tag{23}$$

Since ρ and σ are fixed, we can regard $\rho-\sigma$ as a function of n only, say $\rho-\sigma = \theta(n)$. Then (19) and (23) give

$$[n^{1/d}] \leqslant \theta(n) \leqslant n \tag{24}$$

and

$$f(n) \leqslant dn(n+1)e^{\kappa\theta(n)}f\{n-\theta(n)\}. \tag{25}$$

Now let n_0 be a positive integer and write

$$n_{j+1} = n_j-\theta(n_j) \quad (j = 0, 1, 2,...). \tag{26}$$

By (24), $\theta(n)$ is a positive integer if n is, and $n-\theta(n) \geqslant 0$. Hence there exists a positive integer $J = J(n_0)$ such that

$$n_0 > n_1 > ... > n_J = 0. \tag{27}$$

By (25), (26), and (27),

$$f(n_j) \leqslant dn_0(n_0+1)e^{\kappa(n_j-n_{j+1})}f(n_{j+1}); \tag{28}$$

and hence, by repeated application of (28),

$$f(n_0) \leqslant \{dn_0(n_0+1)\}^{J(n_0)}e^{\kappa n_0}. \tag{29}$$

The inequality

$$J(n) \leqslant \frac{2d}{(d-1)}n^{(d-1)/d} \tag{30}$$

is trivially true for $n = 1$ since $J(1) = 1$. We shall establish the truth of (30) for all n by induction. For suppose that (30) holds for all $n < \nu$, where $\nu \geqslant 2$. Then, by (24),

$$\begin{aligned} J(\nu) = 1+J\{\nu-\theta(\nu)\} &\leqslant 1+\frac{2d}{(d-1)}\{\nu-\theta(\nu)\}^{(d-1)/d} \\ &\leqslant 1+\frac{2d}{(d-1)}\{\nu-[\nu^{1/d}]\}^{(d-1)/d} \\ &= 1+\frac{2d}{(d-1)}\nu^{(d-1)/d}\left\{1-\nu^{-(d-1)/d}\frac{[\nu^{1/d}]}{\nu^{1/d}}\right\}^{(d-1)/d} \\ &= 1+\frac{2d}{(d-1)}\nu^{(d-1)/d}-\frac{2[\nu^{1/d}]}{\nu^{1/d}}\left\{1-\epsilon\nu^{-(d-1)/d}\frac{[\nu^{1/d}]}{\nu^{1/d}}\right\}^{-1/d} \end{aligned} \tag{31}$$

for some ϵ satisfying $0 < \epsilon < 1$. Hence

$$J(\nu) \leqslant 1+\frac{2d}{(d-1)}\nu^{(d-1)/d}-\frac{2[\nu^{1/d}]}{\nu^{1/d}} \leqslant \frac{2d}{(d-1)}\nu^{(d-1)/d}, \tag{32}$$

and (32) shows that (30) is then true for all $n < \nu+1$. This completes the proof of (30).

Thus (3) follows from (1), (29), and (30). The inequalities leading

from (25) to (3) could be sharpened a little at the expense of additional complications; but the net improvement in (3) would not be very great.

REFERENCES

1. J. M. Hammersley, 'Percolation processes II. The connective constant', *Proc. Cambridge Phil. Soc.* 53 (1957) 642–5.
2. E. Hille, *Functional analysis and semi-groups* (1948). American Math. Soc. Colloq. Publ., vol. xxxi.
3. M. F. Sykes, 'Some counting theorems in the theory of the Ising model and the excluded volume problem', *J. Math. Phys.* 2 (1961) 52–62.
4. C. Domb and M. F. Sykes, 'Use of series expansions for the Ising model susceptibility and excluded volume problem', ibid. 63–67.
5. J. M. Hammersley, 'The number of polygons on a lattice' *Proc. Cambridge Phil. Soc.* 57 (1961) 516–23.
6. B. J. Hiley and M. F. Sykes, 'Probability of initial ring closure in the restricted random-walk model of a macromolecule', *J. Chem. Phys.* 34 (1961) 1531–7.
7. M. P. O'Flaherty, 'Computations for the excluded volume problem in connection with the conformation of polymer molecules', *Stanford Univ. Appl. Math. and Statist. Lab. Tech. Rep.* No. 11 (3 July 1961).

A MULTIDIMENSIONAL GENERALIZATION OF THE INVERSE SINE FUNCTION

By H. RUBEN (*New York and Sheffield*)

[Received 21 November 1960]

1. Introduction and summary

THE purpose of this paper is to bring to light a class of functions, the h-functions of equation (17), defined therein as integrals over unit hypercubes, which appear to constitute natural generalizations of the inverse sine function: in the particular case when the unit 'cube' is a square, the corresponding integral does, in fact, serve to define the latter function. The measure of a regular simplex in hyperspherical space may be expressed as a linear combination of such integrals. This representation is achieved in equation (16). In summary, equations (16) and (17) constitute the main results of this paper.

The surprising simplicity of the integrals gives some grounds for hope of further progress in the difficult problem of evaluating the measures of regular hyperspherical simplices. It is hoped subsequently to extend the integrals to cover the case of arbitrary simplices in hyperspherical space.†

2. Reduction of the measure of regular hyperspherical simplices to integrals over hypercubes

Denote the $(n-1)$-dimensional normed measure of a simplex on the surface of an n-sphere with common dihedral angle θ by $u_n(\theta)$, so that the u-function is related to the Schläfli F-function in (**10**) by

$$u_n(\theta) = \frac{n!}{2^n} F_n(\tfrac{1}{2}\theta).$$

† For some previous work on this problem, in addition to Schläfli's fundamental pioneering papers (**10**), (**11**), see Poincaré (**5**), Dehn (**4**), Coxeter (**2**), (**3**), Ruben (**9**), Rogers (**6**), van der Vaart (**12**), (**13**), and Böhm (**1**). In particular, van der Vaart has derived a formula for the measure of an $(n-1)$-dimensional spherical orthoscheme which bears a superficial resemblance to the main formula [equation (16)] of this paper relating to the measure of an $(n-1)$-dimensional regular spherical simplex. (Such a simplex may be decomposed into $n!$ congruent orthoschemes with all non-right angles but one equal to $\tfrac{1}{2}\pi$.) However, van der Vaart's formula appropriate to the regular simplex appears to be far more difficult to handle.

It has been shown previously [Ruben (**7**, **8**)] that, if $\theta > \frac{1}{2}\pi$, then this measure can be expressed in the form of a univariate integral†

$$u_n\left(\pi - \arc\cos\frac{1}{1+\alpha}\right) = \left(\frac{\alpha}{2\pi}\right)^{\frac{1}{2}} \int_{-\infty}^{\xi} e^{-\frac{1}{2}\alpha\xi^2}\{\Phi(\xi)\}^n \, d\,\xi,$$

where α is real and positive and

$$\Phi(\xi) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\xi} e^{-\frac{1}{2}\eta^2}\, d\eta.$$

On replacing $1/(1+\alpha)$ by x, the normed measure of an $(n-1)$-dimensional regular spherical simplex with common *obtuse* dihedral angle $\pi - \arc\cos x$ is obtained as

$$u_n(\pi - \arc\cos x) = \left(\frac{2\pi x}{1-x}\right)^{-\frac{1}{2}} \int_{-\infty}^{\infty} \exp\left(-\frac{1-x}{2x}\xi^2\right)\{\Phi(\xi)\}^n \, d\xi \qquad (0 < x < 1). \quad (1)$$

Equation (1)‡ provides the starting-point for the present discussion.

Define

$$w_m(x) = (2\pi)^{\frac{1}{2}m}\left(\tfrac{1}{2}\pi\frac{x}{1-x}\right)^{-\frac{1}{2}} \int_{0}^{\infty} \exp\left(-\frac{1-x}{2x}\xi^2\right)\{\psi(\xi)\}^m \, d\xi \qquad (m = 0, 2, 4, \ldots;\ 0 < x < 1), \quad (2)$$

where

$$\psi(\xi) = (2\pi)^{-\frac{1}{2}} \int_{0}^{\xi} e^{-\frac{1}{2}\eta^2}\, d\eta.$$

(Note in passing that $w_0(x) \equiv 1$ when $0 < x < 1$.) The right-hand member of (1) can be expressed in terms of the w-functions. To achieve this, represent the integral in (1) as the sum of two integrals with ranges $\xi < 0$ and $\xi \geqslant 0$, respectively. On substituting $-\xi$ for ξ in the first of these two integrals and observing that

$$\Phi(\xi) = \tfrac{1}{2} + \psi(\xi), \qquad \psi(-\xi) = -\psi(\xi),$$

† This result is actually a special case of a somewhat more general result in (**7**) and (**8**). In those papers, the normed measure of a *skew-regular* spherical simplex was expressed as a univariate integral, where the latter simplex is generated from a regular simplex as base by the projection of a (possibly empty) subset of vertices of the base simplex with respect to the centre of the sphere on whose surface the base simplex is located.

‡ It is convenient to define $u_0(\theta) \equiv 1$. Equation (1) then holds for all non-negative integral n.

for all real ξ, we obtain

$$u_n(\pi-\arc\cos x)$$

$$=\left(\frac{2\pi x}{1-x}\right)^{-\frac{1}{2}}\int\limits_0^\infty \exp\left(-\frac{1-x}{2x}\xi^2\right)\{[\tfrac{1}{2}-\psi(\xi)]^n+[\tfrac{1}{2}+\psi(\xi)]^n\}\,d\xi$$

$$=\left(\frac{2\pi x}{1-x}\right)^{-\frac{1}{2}}\int\limits_0^\infty 2\exp\left(-\frac{1-x}{2x}\xi^2\right)\sum_{i=0}^{[\frac{1}{2}n]}\binom{n}{2i}(\tfrac{1}{2})^{n-2i}\{\psi(\xi)\}^{2i}\,d\xi.$$

Therefore

$$u_n(\pi-\arc\cos x)=(\tfrac{1}{2})^n\sum_{i=0}^{[\frac{1}{2}n]}\binom{n}{2i}\left(\frac{2}{\pi}\right)^i w_{2i}(x)\quad(n=0,1,2,...;\ 0<x<1). \tag{3}$$

For the purpose of expressing the w-functions in an alternative and more extended form we now examine a more inclusive class of integrals than that given in (2). At the same time this will serve to generalize the relationship (3). Define

$$w_m^*(x;t_1,...,t_m)=(2\pi)^{\frac{1}{2}m}\left(\tfrac{1}{2}\pi\frac{x}{1-x}\right)^{-\frac{1}{2}}\int\limits_0^\infty \exp\left(-\frac{1-x}{2x}\xi^2\right)\prod_1^m\psi(t_i^2\xi)\,d\xi$$

$$(m=2,4,...;\ 0<x<1), \tag{4}$$

for real t_i, so that, in particular,

$$w_m^*(x;1,...,1)\equiv w_m(x). \tag{5}$$

On differentiating in (4) with respect to the t_i (justification of the differentiation under the integral sign is trivial), we find

$$\frac{\partial^m}{\partial t_1...\partial t_m}w_m^*(x;t_1,...,t_m)$$

$$=(2\pi)^{\frac{1}{2}m}\left(\tfrac{1}{2}\pi\frac{x}{1-x}\right)^{-\frac{1}{2}}\int\limits_0^\infty \exp\left(-\frac{1-x}{2x}\xi^2\right)\xi^m\prod_1^m(2\pi)^{-\frac{1}{2}}\exp(-\tfrac{1}{2}t_i^2\xi^2)\,d\xi$$

$$=\left(\tfrac{1}{2}\pi\frac{x}{1-x}\right)^{-\frac{1}{2}}\int\limits_0^\infty \xi^m\exp\left[-\tfrac{1}{2}\left(\frac{1-x}{x}+\sum_1^m t_i^2\right)\xi^2\right]d\xi$$

$$=1.3...(m-1)\left(\frac{x}{1-x}\right)^{\frac{1}{2}m}\left(1+\frac{x}{1-x}\sum_1^m t_i^2\right)^{-\frac{1}{2}(m+1)}. \tag{6}$$

The complete solution of (6) is

$$w_m^*(x;\mathbf{t})=1.3...(m-1)\left(\frac{x}{1-x}\right)^{\frac{1}{2}m}\int\limits_0^{t_1}...\int\limits_0^{t_m}\left(1+\frac{x}{1-x}\sum_1^m z_i^2\right)^{-\frac{1}{2}(m+1)}dz_1...dz_m+$$

$$+\sum_1^m C_{m;i}(x;\mathbf{t}^{(i)}),$$

where, for convenience, the vectors $(t_1,\ldots,t_m)$, $(t_1,\ldots,t_{i-1},t_{i+1},\ldots,t_m)$ have been replaced by $\mathbf{t}$ and $\mathbf{t}^{(i)}$, respectively, and $w_m^*(x;\,t_1,\ldots,t_m)$ by $w_m^*(x;\,\mathbf{t})$. Equation (7) follows from the fact that $\sum C_{m;i}(x;\,\mathbf{t}^{(i)})$ is the general solution of the differential equation

$$\partial^m w_m^*/\partial t_1\ldots\partial t_m = 0.$$

We now show that $\sum C_{m;i}$ is identically zero. Observe that, by (4),

$$w_m^*(x;\,t_1,\ldots,t_{i-1},0,t_{i+1},\ldots,t_m) = 0 \quad (i = 1,\ldots,m),$$

and set $t_i = 0$ in (7) for a fixed i. Then (7) reduces to

$$C_{m;i}(x;\,\mathbf{t}^{(i)})+\Big[\sum_{j\neq i} C_{m;j}(x;\,\mathbf{t}^{(j)})\Big]_{t_i=0} = 0:$$

that is, $C_{m;i}(x;\,\mathbf{t}^{(i)})$ can be expressed as a sum of functions each involving $m-2$ t-arguments selected from the set $(t_1,\ldots,t_{i-1},t_{i+1},\ldots,t_m)$. Hence $\sum C_{m;i}$ can be expressed in the form

$$\sum_i C_{m;i}(x;\,\mathbf{t}^{(i)}) = \sum_i \sum_j C_{m;i,j}(x;\,\mathbf{t}^{(i,j)}),$$

where $\mathbf{t}^{(i,j)}$ denotes the vector

$$(t_1,\ldots,t_{i-1},t_{i+1},\ldots,t_{j-1},t_{j+1},\ldots,t_m)$$

and $C_{m;i,j}(x;\,\mathbf{t}^{(i,j)})$ represents a function independent of t_i and t_j, summation being extended over all pairs (i,j), with $i<j$, selected from the set $\{1,2,\ldots,m\}$. Similarly, by setting fixed pairs $(t_i,t_j) = (0,0)$ in (7) we obtain

$$\sum_i C_{m;i}(x;\,\mathbf{t}^{(i)}) = \sum_i \sum_j \sum_k C_{m;i,j,k}(x;\,\mathbf{t}^{(i,j,k)}),$$

where $\mathbf{t}^{(i,j,k)}$ denotes the vector

$$(t_1,\ldots,t_{i-1},t_{i+1},\ldots,t_{j-1},t_{j+1},\ldots,t_{k-1},t_{k+1},\ldots,t_m),$$

and $C_{m;i,j,k}(x;\,\mathbf{t}^{(i,j,k)})$ represents a function independent of t_i, t_j, t_k, summation being extended over all triples (i,j,k) from $\{1,2,\ldots,m\}$ with $i<j<k$. Proceeding in this manner, we eventually obtain

$$\sum_i C_{m;i}(x;\,\mathbf{t}^{(i)}) = C_{m;2,3,\ldots,m}(x;\,t_1)+\ldots+C_{m;1,2,\ldots,m-1}(x;\,t_m),$$

and on setting each component of $(m-1)$-tuples of the t_i simultaneously zero it is deduced that $\sum_i C_{m;i}(x;\,\mathbf{t}^{(i)})$ is independent of the t_i. Finally, on setting $(t_1,\ldots,t_m) = (0,\ldots,0)$ the required result

$$\sum_{i=1}^m C_{m;i}(x;\,\mathbf{t}^{(i)}) \equiv 0 \tag{8}$$

is established.

On using (8) in (7), setting $t_1 = 1,\ldots, t_m = 1$ and referring to (5), we express $w_m(x)$ as an integral over a unit m-dimensional cube in the form

$$w_m(x) = 1.3\ldots(m-1)\left(\frac{x}{1-x}\right)^{\frac{1}{2}m} \int\limits_0^1 \ldots \int\limits_0^1 \left(1+\frac{x}{1-x}\sum_1^m z_i^2\right)^{-\frac{1}{2}(m+1)} dz_1\ldots dz_m$$
$$(m = 2, 4,\ldots;\ 0 < x < 1). \quad (9)$$

Equation (9) suggests an obvious extension of the domain of the w-functions since the right-hand member of (9) is well-defined for all x in the closed interval $[-1/(m-1), 1]$.† In fact, define

$$\left.\begin{aligned} h_m(x) &= 1.3\ldots(m-1)\left(\frac{x}{1-x}\right)^{\frac{1}{2}m} \int\limits_0^1 \ldots \int\limits_0^1 \left(1+\frac{x}{1-x}\sum_1^m z_i^2\right)^{-\frac{1}{2}(m+1)} dz_1\ldots dz_m \\ &\qquad (m = 2, 4,\ldots;\ -1/(m-1) \leqslant x \leqslant 1) \\ h_0(x) &\equiv 1 \end{aligned}\right\}, \quad (10)$$

so that $h_m(x) = w_m(x)$ for $0 < x < 1$. Note further that $h_m(0) = 0$.

Equation (3) was shown previously to hold for $0 < x < 1$. By analytic continuation it clearly holds for all values of x for which $u_n(x)$ is defined, with $w_{2i}(x)$ replaced by $h_{2i}(x)$. Accordingly,

$$u_n(\pi-\operatorname{arc}\cos x) = (\tfrac{1}{2})^n \sum_{i=0}^{[\frac{1}{2}n]} \binom{n}{2i}\left(\frac{2}{\pi}\right)^i h_{2i}(x)$$
$$(n = 0, 1, 2,\ldots;\ -1/(n-1) \leqslant x \leqslant 1). \quad (11)$$

Equation (11) now permits of a further useful extension of the domain of the original w-functions. On inverting (11) to solve for the h's in terms of the u's, we find

$$h_n(x) = (2\pi)^{\frac{1}{2}n} \sum_{i=0}^{n} \binom{n}{i}(-\tfrac{1}{2})^i u_{n-i}(\pi-\operatorname{arc}\cos x)$$
$$(n = 0, 2, 4,\ldots;\ -1/(n-1) \leqslant x \leqslant 1). \quad (12)$$

Now it is well known that $u_{2k+1}(x)$ is linearly related to $u_0(x), u_2(x),\ldots, u_{2k}(x)$ for $k = 0, 1, 2,\ldots$ [Schläfli (**10**)]. This implies, of course, that $u_{2k+1}(x)$ is a linear function of $u_0(x), u_1(x),\ldots, u_{2k}(x)$ for $k = 0, 1, 2,\ldots$. The linear function in question is readily shown from Schläfli's relationship to be

$$\sum_{i=0}^{2k+1} \binom{2k+1}{i}(-\tfrac{1}{2})^i u_{2k+1-i}(\pi-\operatorname{arc}\cos x) = 0$$
$$(k = 0, 1, 2,\ldots;\ -\tfrac{1}{2}k^{-1} \leqslant x \leqslant 1). \quad (13)$$

† We take $w_m(1) \equiv w_m(1-) < \infty$.

A comparison of (12) and (13) shows that the series in the two equations are identical when n is replaced by $2k+1$ in the former equation ($k = 0, 1, 2,...$). This suggests defining

$$h_{2k+1}(x) \equiv 0 \quad (k = 0, 1, 2,...). \tag{14}$$

Equation (12) is now replaced by

$$h_n(x) = (2\pi)^{\frac{1}{2}n} \sum_{i=0}^{n} \binom{n}{i} (-\tfrac{1}{2})^i u_{n-i}(\pi - \text{arc}\cos x)$$
$$(n = 0, 1, 2,...;\ -1/(n-1) \leqslant x \leqslant 1), \tag{15}$$

while (11) is replaced by

$$u_n(\pi - \text{arc}\cos x) = (\tfrac{1}{2})^n \sum_{i=0}^{n} \binom{n}{i} \left(\frac{2}{\pi}\right)^{\frac{1}{2}i} h_i(x)$$
$$(n = 0, 1, 2,...;\ -1/(n-1) \leqslant x \leqslant 1), \tag{16}$$

where

$$\left.\begin{aligned} h_0(x) &= 1, \\ h_m(x) &= 1.3...(m-1)\left(\frac{x}{1-x}\right)^{\frac{1}{2}m} \int_0^1 ... \int_0^1 \left(1 + \frac{x}{1-x}\sum_1^m z_i^2\right)^{-\frac{1}{2}(m+1)} dz_1...dz_m \\ & \qquad\qquad (m = 2, 4,...;\ -1/(m-1) \leqslant x \leqslant 1) \\ h_m(x) &= 0 \qquad\qquad (m = 1, 3,...) \end{aligned}\right\}. \tag{17}$$

(It should be remarked that $h_m(1) = (\frac{1}{2}\pi)^{\frac{1}{2}m}$ for $m = 0,\ 2,...$. This follows from (12) since $u_n(\pi) = \frac{1}{2}$.)

For $m = 2$ in (17), we find

$$\begin{aligned} h_2(x) &= \frac{x}{1-x} \int_0^1 \int_0^1 \left\{1 + \frac{x}{1-x}(z_1^2 + z_2^2)\right\}^{-\frac{3}{2}} dz_1\, dz_2 \\ &= \frac{2x}{1-x} \int_0^{\frac{1}{4}\pi} \int_0^{\sec\phi} \left(1 + \frac{x}{1-x} r^2\right)^{-\frac{3}{2}} r\, dr d\phi \\ &= 2 \int_0^{\frac{1}{4}\pi} \left\{1 - \left(1 + \frac{x}{1-x}\sec^2\phi\right)^{-\frac{1}{2}}\right\} d\phi \\ &= 2 \int_0^{\frac{1}{4}\pi} \left[1 - \frac{(1-x)^{\frac{1}{2}}\cos\phi}{\{1-(1-x)\sin^2\phi\}^{\frac{1}{2}}}\right] d\phi \\ &= \tfrac{1}{2}\pi - 2\,\text{arc}\sin\left(\frac{1-x}{2}\right)^{\frac{1}{2}} \\ &= \text{arc}\sin x. \end{aligned} \tag{18}$$

In this sense $h_{2k}(x)$, for $k > 1$, is a generalization of the inverse sine-function (as indicated in the introductory section).

We conclude by remarking that the h-functions may be expressed as power series in $x/(1-x)$, provided that $-1/(2k-1) < x < 1/(2k+1)$. Thus, on expanding the integrand in (17) in powers of $x/(1-x)$ and integrating term by term, we have

$$h_{2k}(x) = 1.3...(2k-1)\left(\frac{x}{1-x}\right)^k \int_0^1 ... \int_0^1 \sum_{s=0}^{\infty} \frac{(-)^s}{s!}\left(\frac{2k+1}{2}\right)_{(s)} \times$$

$$\times \left(\sum_{i=1}^{2k} t_i^2\right)^s \left(\frac{x}{1-x}\right)^s dt_1...dt_{2k}$$

$$= 1.3...(2k-1)\left(\frac{x}{1-x}\right)^k \sum_{s=0}^{\infty} c_{2k;s}\left(\frac{x}{1-x}\right)^s$$

$$(k = 1,2,...;\ -1/(2k-1) < x < 1/(2k+1)), \quad (19)$$

where

$$c_{2k;s} = \frac{(-1)^s}{s!}\left(\frac{2k+1}{2}\right)_{(s)} \int_0^1 ... \int_0^1 \left(\sum_1^{2k} t_i^2\right)^s dt_1...dt_{2k}$$

$$= (-1)^s\left(\frac{2k+1}{2}\right)_{(s)} \int_0^1 ... \int_0^1 \sum_{j_1+...+j_{2k}=s} \frac{t_1^{2j_1}...t_{2k}^{2j_{2k}}}{j_1!\,...\,j_{2k}!}\, dt_1 ... dt_{2k}$$

$$= (-1)^s\left(\frac{2k+1}{2}\right)_{(s)} \cdot \sum_{j_1+...+j_{2k}=s} \frac{1}{j_1!...j_{2k}!(2j_1+1)...(2j_{2k}+1)}$$

$$(s = 0, 1, 2,...),$$

and $\quad q_{(0)} = 1, \quad q_{(s)} = q(q+1)...(q+s-1) \quad (s = 1, 2,...).$

Equation (19) is a suitable formula for the evaluation of $u_n(\theta)$)when $\cos\theta > -1/(2k+1)$, i.e. when the simplex is not too 'large'.

REFERENCES

1. Johannes Böhm, 'Untersuchung des Simplexinhaltes in Räumen konstanter beliebiger Dimension', *J. reine angew. Math.* 202 (1959) 16–51.
2. H. S. M. Coxeter, 'The functions of Schläfli and Lobachewski', *Quart. J. of Math.* (Oxford) 6 (1935) 13–29.
3. —— *Regular polytopes* (London, 1948).
4. M. Dehn, 'Die Eulersche Formel im Zusammenhang mit dem Inhalt in der Nichteuklidischen Geometrie', *Math. Ann.* 61 (1905) 561–74.
5. H. Poincaré, 'Sur la généralisation d'un théorème élémentaire de géométrie', *C.R. Acad. Sci. Paris* 140 (1905) 113–7.
6. C. A. Rogers, 'An asymptotic expansion for certain Schläfli functions', *J. London Math. Soc.* 36 (1961) 77–80.

7. H. Ruben, 'On the moments of order statistics in samples from normal populations', *Biometrika* 41 (1954) 200–27.

8. —— 'On the geometrical moments of skew-regular simplices in hyperspherical space, with some applications in geometry and mathematical statistics', *Acta Math.* 103 (1960) 1–23.

9. —— 'A power series expansion for a class of Schläfli functions', *J. London Math. Soc.* 36 (1961) 69–77.

10. L. Schläfli, 'On the multiple integral $\int^n dxdy...dz$ whose limits are

$$p_1 = a_1x+b_1y+...+h_1z > 0, p_2 > 0,..., p_n > 0, \text{ and } x^2+y^2+...+z^2 < 1\text{'}.$$

Quart. J. of Math. 2 (1858) 269–301; 3 (1860) 54–68, 97–108.

11. —— *Neue Denkschriften der allgemeinen schweizerischen Gesellschaft für die gesamten Naturwissenschaften* (Berne, 1901).

12. H. R. van der Vaart, 'The content of certain spherical polyhedra for any number of dimensions', *Experientia* (1953) 88–89.

13. —— 'The content of some classes of non-Euclidean polyhedra for any number of dimensions with several applications. I and II', *Nederlandse Akad. Wetensch. Proc.* A, 58 (1955) 200–21.

[*This research was supported in part by the Office of Naval Research.*]

APPROXIMATE LOCATION OF THE ZEROS OF GENERALIZED BESSEL POLYNOMIALS

By P. J. McCARTHY (*University of Kansas*)

[Received 14 January 1961]

In this note we consider the generalized Bessel polynomials in Al-Salam's notation (**1**)

$$Y_n^{(\alpha)}(z) = \sum_{k=0}^{n} \binom{n}{k} (n+\alpha+1)_k (\tfrac{1}{2}z)^k,$$

where $(\beta)_k = \Gamma(\beta+k)/\Gamma(\beta)$. In the notation of Krall and Frink (**2**) we have

$$Y_n^{(\alpha)}(z) = y_n(z, \alpha+2, 2).$$

Al-Salam has shown [(**1**) Theorem 12.2] that, if $\alpha > 0$, then all the zeros of $Y_n^{(\alpha)}(z)$ are inside the circle $|z| = 1$. On the other hand, Nassif has shown in (**3**) that, if $\alpha = 0$, in which case we have the ordinary Bessel polynomials, and, if $n > 1$, then all of the zeros of $Y_n^{(\alpha)}(z)$ are inside or on the circle $|z| = \{(n-1)/(2n-1)\}^{\frac{1}{2}}$. In this note I shall obtain two results concerning the location of the zeros of $Y_n^{(\alpha)}(z)$.

THEOREM 1. *For fixed n and sufficiently large positive α, all the zeros of $Y_n^{(\alpha)}(z)$ are inside the left half of the circle $|z| = 1$. As $\alpha \to \infty$, the zeros of $Y_n^{(\alpha)}(z)$ approach zero.*

THEOREM 2. *For fixed $\alpha \geqslant 0$ and sufficiently large n, all the zeros of $Y_n^{(\alpha)}(z)$ are inside the circle $|z| = \{(n-1)/(2n+\alpha-1)\}^{\frac{1}{2}}$.*

Theorem 1 is a consequence of a general result due to Parodi [(**4**) 150–5]. This result concerns the location of the zeros of the polynomials in a sequence of polynomials which satisfy a three-term recurrence relation. For the sequence of Bessel polynomials $\{Y_n^{(\alpha)}(z)\}$ we have the recurrence relation [(**2**) 111]

$$(n+\alpha)(2n+\alpha-2)Y_n^{(\alpha)}(z) - \\ -(2n+\alpha-1)\{\alpha+\tfrac{1}{2}(2n+\alpha)(2n+\alpha-2)x\}Y_{n-1}^{(\alpha)}(z) - \\ -(n-1)(2n+\alpha)Y_{n-2}^{(\alpha)}(z) = 0.$$

Then, in Parodi's notation,

$$f_1(n) = (n+\alpha)(2n+\alpha-2), \qquad f_2(n) = -\alpha(2n+\alpha-1),$$
$$f_3(n) = \tfrac{1}{2}(2n+\alpha)(2n+\alpha-1)(2n+\alpha-2),$$
$$f_4(n) = -(n-1)(2n+\alpha).$$

I shall not use Parodi's complete result, but only the fact that, for α sufficiently large and positive, the zeros of $Y_n^{(\alpha)}(z)$ are in the union of the circular regions given by

$$\left|z+\frac{2\alpha}{(2n+\alpha)(2n+\alpha-2)}\right| \leqslant \left|\frac{2(n-1)}{(2n+\alpha-1)(2n+\alpha-2)}\right|,$$

$$\left|z+\frac{2\alpha}{(2k+\alpha)(2k+\alpha-2)}\right| \leqslant \left|\frac{2\alpha}{(2k+\alpha)(2k+\alpha-2)}\right|,$$

for $k = 2, 3,\ldots, n-1$, and

$$\left|z+\frac{2}{2+\alpha}\right| \leqslant \left|\frac{2}{2+\alpha}\right|.$$

Since $\alpha > 0$, then all the zeros are inside the circle $|z| = 1$. Each of the above circles has its centre on the left half of the real axis, and only the first could possibly contain any points of the open right half-plane. Moreover, if $\alpha \geqslant \frac{1}{2}\{-n+\sqrt{(9n^2-8n)}\}$, even the first circle does not intersect the open right half-plane. Thus, noting that $Y_n^{(\alpha)}(0) \neq 0$, we see that, for all such α, all the zeros of $Y_n^{(\alpha)}(z)$ are inside the left half of the circle $|z| = 1$.

We also note that both the centres and the radii of the above circular regions approach zero as $\alpha \to \infty$. Finally, we note that the zeros of $Y_n^{(\alpha)}(z)$, for large α, are bounded away from the imaginary axis except near zero.

I shall give only an outline of the proof of Theorem 2 since it is similar to the proof given by Nassif in the case $\alpha = 0$. I shall indicate, however, why it is that we must choose n large. We assume $n \geqslant 3$ and set

$$c_k = 2^{-k}\binom{n}{k}(n+\alpha+1)_k.$$

Following Nassif, we consider the polynomial

$$R_n(x) = c_n x^n - c_{n-1} x^{n-1} + c_{n-2} x^{n-2} - c_{n-3} x^{n-3}$$

$$= (n-1)^{-1} c_{n-2} x^{n-3}\left\{\frac{(2n+\alpha)(2n+\alpha-1)}{2n} x^3 - \right.$$

$$\left. -(2n+\alpha-1)x^2+(n-1)x-\frac{2(n-1)(n-2)}{3(2n+\alpha-2)}\right\}.$$

Let $f(x)$ be the polynomial in the brackets and set

$$r = \{(2n+\alpha-1)/(n-1)\}^{\frac{1}{2}}.$$

We wish to show that, for n sufficiently large, $f(x)$ has exactly one real zero which lies between 0 and $1/r$. We have $f(0) < 0$ and we consider

$f'(x)$. The discriminant of $f'(x)$ is

$$n^{-1}\{-8n^3+(20-8\alpha)n^2-(2\alpha^2-22\alpha+8)n+6\alpha(\alpha-1)\},$$

which is negative for n sufficiently large. Hence $f(x)$, for n sufficiently large, has at most one real zero. Since, with n sufficiently large, $f(1/r) > 0, f(x)$ has exactly one zero between 0 and $1/r$. Thus $R_n(x) > 0$ for n sufficiently large and $x \geqslant 1/r$. We may now assume $n \geqslant 4$. Then, if $x \geqslant 1/r$ and $0 \leqslant k \leqslant n-4$, we have $c_k x^k < c_{k+1} x^{k+1}$ if only we choose n large enough. Then it follows, as in Nassif's proof, that $Y_n^{(\alpha)}(-x) > 0$ for $x \geqslant 1/r$. Hence, if $Y_n^{(\alpha)}(z)$ has a real zero x, we must have $-1/r < x < 0$; this, of course, with n sufficiently large.

Now let
$$g(z) = (rz)^n Y_n^{(\alpha)}(1/rz) = \sum_{k=0}^{n} a_k z^k,$$
where $a_k = c_{n-k} r^k$. We have

$$\begin{aligned} ra_0 &= 2^{-n+1}(n+\tfrac{1}{2}\alpha)(n+\alpha+1)_{n-1} r \\ &> 2^{-n+1}n(n+\alpha+1)_{n-1} r = a_1 > 0. \end{aligned}$$

Also, if $1 \leqslant k \leqslant n-1$,

$$\frac{ra_{k+1}}{a_k} = \frac{2(n-k)r^2}{(k+1)(2n+\alpha-k)} < 1,$$

if we choose n large enough. Thus, for n sufficiently large, we have

$$a_k > ra_{k+1} > 0, \qquad 1 \leqslant k \leqslant n-1.$$

We note that in each case where we are required to choose n large it is because we wish some polynomial in n to be positive or to be negative for all the n under consideration. We can now choose n so large that all of the desired inequalities hold. Then we can proceed as did Nassif [(3) 411] to show that $g(z)$ has at most one zero inside or on the unit circle. Thus $Y_n^{(\alpha)}(z)$ has at most one zero outside or on the circle $|z| = 1/r$, which must therefore be a real zero. But all the real zeros of $Y_n^{(\alpha)}(z)$ are inside this circle, and so all of the zeros of $Y_n^{(\alpha)}(z)$ must be inside this circle.

REFERENCES

1. W. A. Al-Salam, 'Some functions related to the Bessel polynomials', *Duke Math. J.* 26 (1959) 519–39.
2. H. L. Krall and O. Frink, 'A new class of orthogonal polynomials: the Bessel polynomials', *Trans. American Math. Soc.* 65 (1949) 100–15.
3. M. Nassif, 'Note on the Bessel polynomials', ibid. 77 (1954) 408–12.
4. M. Parodi, *La localisation des values caractéristiques des matrices et ses applications* (Paris, 1959).

[*This work was sponsored by the National Science Foundation.*]

ON THE COHOMOLOGY OF AN EILENBERG–MACLANE SPACE

By W. D. BARCUS (*Brown University, U.S.A.*)

[Received 23 June 1961]

1. Introduction

LET π be a finitely-generated Abelian group. Cartan's calculation (**2**) of the cohomology $H^*(\pi, n; Z_p)$ of an Eilenberg–MacLane space determines the local structure as a module over the Steenrod algebra. We shall compute the module structure in the large in dimensions less than pn (Propositions 4.1, 5.6, and 5.7), including a minimal set of generators in dimensions less than $3n$ (Propositions 4.5, 5.11, and 5.12). The latter is the dimension range most useful for applications (cf. § 6).

In revising this paper I am indebted to W. Browder for suggesting the use of the co-algebra EA instead of just a filtration, and to D. Jonah for simplifying Lemma 2.2.

2. Modules over Hopf algebras

If B is any graded associative algebra over a field Λ, with a unit e, then the terms 'module' (or 'B-module', if B is not clear from context) and 'homomorphism' are to mean 'graded unitary left B-module' and 'graded B-homomorphism'. The term 'vector-space' is to mean 'graded Λ-vectorspace', and 'element' is to mean 'homogeneous element'. An algebra will often be considered as a module over itself under left multiplication. The *r-truncation* of a module M is the quotient of M by the submodule $\sum M_i\,(i \geqslant r)$. If a module M is regraded by raising the dimension of each element by a fixed integer n, we shall denote the resulting module by M^n.

Let A be a connected Hopf algebra over a field Λ, with associative multiplication $\phi\colon A \otimes A \to A$ and associative and commutative diagonal $\psi\colon A \to A \otimes A$ [for terminology cf. (**4**) or (**5**)]. If M and M' are A-modules, then $M \otimes M'$ is to be the module obtained by allowing A to operate through the diagonal, i.e. if $\psi a = \sum a_i \otimes a_i'$, then

$$a(m \otimes m') = \sum (-1)^{q_i} a_i m \otimes a_i' m',$$

where $q_i = \dim(a_i')\dim(m)$. Similarly A is to operate on a q-fold tensor product through

$$\psi^q = (\psi \otimes 1 \otimes \dots \otimes 1)\dots(\psi \otimes 1)\psi\colon A \to \otimes^q A.$$

Recall that, if M is a vectorspace, then the symmetric algebra

$$SM = \Lambda + M + SP^2M + \ldots + SP^qM + \ldots$$

is obtained from the tensor algebra

$$\otimes M = \Lambda + M + \otimes^2 M + \ldots + \otimes^q M + \ldots$$

as follows. The symmetric group $\mathscr{S}_q$ acts on $\otimes^q M$ by permuting factors with the usual sign convention: thus, if $\sigma \in \mathscr{S}_2$ ($\sigma \neq 1$), then

$$\sigma(m \otimes m') = (-1)^\lambda m' \otimes m, \qquad \lambda = \dim(m')\dim(m),$$

and SP^qM is the quotient of $\otimes^q M$ by the vector subspace generated by all elements $g - \sigma(g)$, for g in $\otimes^q M$, σ in $\mathscr{S}_q$. If M is also an A-module, then the commutativity of ψ implies that SP^qM inherits a module structure from $\otimes^q M$. We write the image in SP^qM of $m_1 \otimes \ldots \otimes m_q$ as $m_1 \ldots m_q$.

Consider the homomorphisms

$$\mu\colon SP^qM \to \otimes^q M, \qquad \nu\colon \otimes^q M \to SP^qM,$$

where ν is the projection and

$$\mu(m_1 \ldots m_q) = \sum \sigma(m_1 \otimes \ldots \otimes m_q) \quad (\sigma \in \mathscr{S}_q).$$

The composition $\nu\mu$ is just multiplication by $q!$. Hence, if q is less than the characteristic of Λ ($= p$ or ∞), then SP^qM is isomorphic to a direct summand of $\otimes^q M$. Now, if M is A-free, then so is $\otimes^q M$: in fact a module basis is given by the elements $\alpha_i \otimes m_{i_2} \otimes \ldots \otimes m_{i_q}$, where α_i runs through an A-basis for M; and for each j, m_{i_j} runs through a Λ-basis. Hence in this case SP^qM is a graded projective module; and, since A is connected, we have by (**3**) (Exposé 15, Proposition 9) the proposition:

PROPOSITION 2.1. *If M is A-free, then SP^qM is A-free provided that q is less than the characteristic of Λ.*

This is not true for $q = p$, as is shown by the example SP^pA. If $a \in A$ is primitive, then $a(e^p) = pe^{p-1}.a = 0$, whereas any set of generators for SP^pA must include a non-zero scalar multiple of e^p.

Consider the special case in which Λ has finite characteristic p (p odd), and A is a divided polynomial algebra on one 'generator' of even dimension r; thus A has a vectorspace basis $e = x_0, x_1, x_2, \ldots$, where x_i has dimension ri,

$$\psi x_i = \sum_{l+m=i} x_l \otimes x_m, \qquad \phi(x_i \otimes x_j) = (i,j)x_{i+j}.$$

Here (i,j) denotes the mod p reduction of the binomial coefficient $(i+j)!/i!j!$. We need a lemma:

LEMMA 2.2. *In $A^n \otimes A^n$ the following relations hold*:

$$e \otimes x_s - (-1)^s x_s \otimes e = -\sum_{r=1}^{s} (-1)^r x_r (e \otimes x_{s-r}), \quad \text{for } s = 1, 2, \ldots.$$

Proof. Let

$$\begin{aligned}
\rho &= \sum_{r=0}^{s} (-1)^r x_r (e \otimes x_{s-r}) \\
&= \sum_{r=0}^{s} (-1)^r \sum_{m=0}^{r} (r-m, s-r) x_m \otimes x_{s-m} \\
&= \sum_{m=0}^{s} \Big[\sum_{r=m}^{s} (-1)^r (s-r, r-m) \Big] x_m \otimes x_{s-m}.
\end{aligned}$$

Setting $q = r-m$, we get

$$\begin{aligned}
\rho &= (-1)^s x_s \otimes e + \sum_{m=0}^{s-1} (-1)^m \Big[\sum_{q=0}^{s-m} (-1)^q (s-m-q, q) \Big] x_m \otimes x_{s-m} \\
&= (-1)^s x_s \otimes e
\end{aligned}$$

since the coefficient inside the square brackets is the alternating sum of the binomial coefficients (reduced mod p), and is therefore zero. Transposing the term $e \otimes x_s$ of ρ, we obtain the desired equality. We have the proposition:

PROPOSITION 2.3. *The set of elements*

$$e \otimes x_s + (-1)^s x_s \otimes e \quad (s = 0, 1, \ldots)$$

forms a basis for $A^n \otimes A^n$ as an A-module.

Proof. Let $N(m)$ be the submodule of $A^n \otimes A^n$ which is generated over A by the elements of dimension not exceeding $\dim x_m$. Since $A^n \otimes A^n$ has a module basis consisting of the elements $e \otimes x_s$ $(s = 0, 1, \ldots)$, $N(m)$ is A-free with basis $e \otimes x_s$ $(0 \leqslant s \leqslant m)$; we must show that it also has as a basis the set $\Gamma(m)$ of elements

$$e \otimes x_s + (-1)^s x_s \otimes e \quad (0 \leqslant s \leqslant m).$$

Since $\Gamma(m)$ contains the correct number of elements for a basis, it suffices to show that the submodule $M(m)$ of $N(m)$ which is generated by $\Gamma(m)$ is equal to $N(m)$; i.e. that $e \otimes x_s \in M(m)$ for $s \leqslant m$. Suppose inductively that $e \otimes x_s \in M(m-1)$ for $s \leqslant m-1$. Then by Lemma 2.2,

$$e \otimes x_m - (-1)^m x_m \otimes e \in N(m-1) = M(m-1) \subset M(m),$$

while $e \otimes x_m + (-1)^m x_m \otimes e \in M(m)$ by definition. Adding, we find that $e \otimes x_m \in M(m)$ since the characteristic of Λ is not 2.

3. A filtration for co-algebras

Recall the following construction from (5), whose notation we follow. Let B be an augmented, graded, associative algebra over a field Λ,

with multiplication ϕ. Let $\bar{B}$ be the kernel of the augmentation, $j\colon \bar{B} \to B$ the inclusion, and set

$$\phi^q = \phi(\phi\otimes 1)...(\phi\otimes 1\otimes ...\otimes 1)\colon \otimes^q B \to B,$$
$$j^q = j\otimes ...\otimes j\colon \otimes^q \bar{B} \to \otimes^q B.$$

Then we filter B by $F_0^0 B = B$, $F_0^1 B = \bar{B}$, and $F_0^q B = \text{Image}\, \phi^q j^q$ for $q \geqslant 2$. Let

$$E_0^q B = F_0^q B/F_0^{q+1}B, \qquad E_0 B = \sum_{q\geqslant 0} E_0^q B.$$

Then ϕ induces a multiplication in $E_0 B$ which turns it into an algebra.

We shall use a dual construction. Let C be an augmented, graded, associative co-algebra over Λ, with diagonal ψ. Let $\bar{C}$ be the cokernel of the augmentation $\eta\colon \Lambda \to C$, $\zeta\colon C \to \bar{C}$ the projection. Filter C by $F^0C = \operatorname{Ker}\zeta$, $F^{q-1}C = \operatorname{Ker}\zeta^q\psi^q\colon C \to \otimes^q C \to \otimes^q \bar{C}$ for $q \geqslant 2$. Set $E^qC = F^qC/F^{q-1}C$, and $EC = \sum_{q\geqslant 0} E^qC$. EC is in fact bigraded: for, if we set $F^{q,r}C = F^qC \cap C_r$, then $F^qC = \sum_{r\geqslant 0} F^{q,r}C$, and $E^qC = \sum E^{q,r}C$, where $E^{q,r}C = F^{q,r}C/F^{q-1,r}C$. An element $\gamma \in E^{q,r}C$ will be called *bihomogeneous*, of filtration degree q and dimension r. We write the image of an element $c \in F^{q,r}C$ as $\bar{c} \in E^{q,r}C$; c will be called a *lift* of $\bar{c}$.

Now, if $c \in F^{q,r}C$, then ψc must have the form $\psi c = \sum c_i \otimes d_i$, with $c_i \in F^{q_i,r_i}C$, $d_i \in F^{s_i,t_i}C$, and $q_i + s_i = q$, $r_i + t_i = r$. Hence ψ induces a homomorphism $\Psi\colon EC \to EC \otimes EC$ by $\Psi\bar{c} = \sum \bar{c}_i \otimes \bar{d}_i$. One verifies readily that EC is a bigraded co-algebra. In fact, if C is of finite type, then $EC = (E_0 C^*)^*$, so that the properties of E follow from those of E_0 [cf. (5)]. In particular, if C is a Hopf algebra, then there is an induced multiplication in EC which turns it into a bigraded Hopf algebra, the dual of $E_0 C^*$.

If $M = \cup F^qM$ is a filtered module, then M^n is to be filtered by $F^qM^n = (F^qM)^n$; if N is also filtered, we set

$$F^q(M\otimes N) = \sum_{r+s=q} F^rM \otimes F^sN.$$

Let A be as in § 2. We shall use the following relationship between $A^n \otimes A^n$ as an A-module and $(EA)^n \otimes (EA)^n$ as an EA-module. Let $\sum \alpha_i(\beta_i \otimes \gamma_i)$ be a bihomogeneous expression in the latter: i.e. α_i, β_i, γ_i of EA are† bihomogeneous of filtration degrees q_i, q_i', q_i'', dimensions r_i, r_i', r_i'' respectively, and

$$q_i + q_i' + q_i'' = q, \qquad r_i + r_i' + r_i'' = r$$

† We need not distinguish between EA and $(EA)^n$ except in the module structure of the tensor product.

are independent of i. Let a_i, b_i, c_i of A be lifts of α_i, β_i, γ_i respectively. Then, if $\sum \alpha_i(\beta_i \otimes \gamma_i) = 0$, in the A-module $A^n \otimes A^n$ we clearly have

$$\sum a_i(b_i \otimes c_i) \equiv 0 \pmod{F^{q-1}(A^n \otimes A^n)}. \tag{3.1}$$

4. The case $\pi = Z_p$

Henceforth p is to be a fixed odd prime; Λ is to be Z_p (the integers mod p); and A is to be the mod p Steenrod algebra [cf. (4) for a definition].

Cartan's theorem (2) on the structure of $H^*(Z_p, n; Z_p)$ implies that there is an epimorphism (of Z_p-algebras and A-modules)

$$\zeta\colon SA^n \to H^*(Z_p, n; Z_p)$$

which induces an isomorphism of pn-truncations. If the elements of the Adem–Cartan basis for A are written St^{I} as usual, then ζ is defined by

$$\zeta(St^{\mathrm{I}_1} \dots St^{\mathrm{I}_q}) = St^{\mathrm{I}_1}b \dots St^{\mathrm{I}_q}b,$$

where b is the basic class. Hence by Proposition 2.1 we have the proposition:

PROPOSITION 4.1. *The pn-truncation of the augmented cohomology $\bar{H}^*(Z_p, n; Z_p)$ is a truncated free A-module.*

In fact, the integer pn may be replaced by $pn+1$ if n is even, and by $pn+p-1$ if n is odd. These restrictions cannot be improved because of the relations $Q_0 \mathscr{P}^{\frac{1}{2}n}b = 0$ for n even, and $\mathscr{P}^{\frac{1}{2}(n+1)}b = 0$ for n odd.

We shall next calculate a basis for SP^2A^n. Recall that Milnor (4) proved that A is a Hopf algebra, and that the dual A^* has the structure as an algebra of a tensor product of polynomial algebras on generators ξ_1, ξ_2,..., where ξ_k has dimension $2p^k-2$, and exterior algebras on generators τ_0, τ_1,..., where τ_k has dimension $2p^k-1$. The diagonal is given by

$$\phi^*(\xi_k) = \sum \xi_{k-i}^{p^i} \otimes \xi_i,$$
$$\phi^*(\tau_k) = \sum \xi_{k-i}^{p^i} \otimes \tau_i + \tau_k \otimes e.$$

Since ψ^* is commutative, $E_0 A^*$ is isomorphic to A^* as an algebra [(5) Proposition 4.18]. The formula for ϕ^* shows that ξ_k and τ_k are primitive modulo elements of filtration greater than one, so that $E_0 A^*$ is isomorphic as a Hopf algebra to the above tensor product of polynomial and exterior algebras. Hence EA is isomorphic as a Hopf algebra to a tensor product of divided polynomial algebras on 'generators' ζ_1, ζ_2,..., and exterior algebras on generators σ_0, σ_1,..., where ζ_k is dual to ξ_k and σ_k to τ_k. The element $Q_{i_1} \dots Q_{i_s} \mathscr{P}^{r_1,\dots,r_t} \in A$ (in Milnor's notation) is thus a lift of $\sigma_{i_1} \dots \sigma_{i_s} \zeta_1^{r_1} \dots \zeta_t^{r_t}$ in the sense of § 3.

We shall consider the A-module $A^n \otimes A^n$ and the EA-module $(EA)^n \otimes (EA)^n$. An element of the former of dimension r (or of the

latter of filtration degree q and dimension r) will be called *reducible* if it can be written as a homogeneous expression $\sum \alpha_i(\beta_i \otimes \gamma_i)$ of dimension r (or as a bihomogeneous expression of filtration degree q and dimension r as in § 3), with $\dim \alpha_i > 0$ for all i. Such an expression will be called a *reduced form* for the element. We prove the lemma:

LEMMA 4.2. *The element*

$$e \otimes Q_{i_1} \dots Q_{i_s} \mathscr{P}^{r_1 \dots r_l} - (-1)^{R+s(n+1)} Q_{i_1} \dots Q_{i_s} \mathscr{P}^{r_1 \dots r_l} \otimes e$$

in $A^n \otimes A^n$ is reducible modulo elements of filtration less than $R+s$, where $R = r_1 + \dots + r_l$.

Proof. For each j, Lemma 2.2 yields a reduced form for the following element of $(EA)^n \otimes (EA)^n$,

$$\zeta_1^{r_1} \dots \zeta_{j-1}^{r_{j-1}} \otimes \sigma_{i_1} \dots \sigma_{i_s} \zeta_j^{r_j} \dots \zeta_l^{r_l} - (-1)^{r_j} \zeta_1^{r_1} \dots \zeta_j^{r_j} \otimes \sigma_{i_1} \dots \sigma_{i_s} \zeta_{j+1}^{r_{j+1}} \dots \zeta_l^{r_l}.$$

Adding the expressions for all j yields a reduced form for

$$e \otimes \sigma_{i_1} \dots \sigma_{i_s} \zeta_1^{r_1} \dots \zeta_l^{r_l} - (-1)^R \zeta_1^{r_1} \dots \zeta_l^{r_l} \otimes \sigma_{i_1} \dots \sigma_{i_s}.$$

If n and s are even, then adding the following equations for $k = 1, \dots, s$ yields a reduced form for the element

$$\zeta_1^{r_1} \dots \zeta_l^{r_l} \otimes \sigma_{i_1} \dots \sigma_{i_s} - (-1)^{s(n+1)} \sigma_{i_1} \dots \sigma_{i_s} \zeta_1^{r_1} \dots \zeta_l^{r_l} \otimes e:$$

$$\sigma_{i_k}(\sigma_{i_{k+1}} \dots \sigma_{i_s} \zeta_1^{r_1} \dots \zeta_l^{r_l} \otimes \sigma_{i_1} \dots \sigma_{i_{k-1}}) = \sigma_{i_k} \dots \sigma_{i_s} \zeta_1^{r_1} \dots \zeta_l^{r_l} \otimes \sigma_{i_1} \dots \sigma_{i_{k-1}} +$$
$$+ (-1)^{s-1} \sigma_{i_{k+1}} \dots \sigma_{i_s} \zeta_1^{r_1} \dots \zeta_l^{r_l} \otimes \sigma_{i_1} \dots \sigma_{i_k};$$

similarly for n or s odd. Hence we obtain a reduced form (which can easily be written explicitly) for

$$e \otimes \sigma_{i_1} \dots \sigma_{i_s} \zeta_1^{r_1} \dots \zeta_l^{r_l} - (-1)^{R+s(n+1)} \sigma_{i_1} \dots \sigma_{i_s} \zeta_1^{r_1} \dots \zeta_l^{r_l} \otimes e.$$

The lemma then follows from (3.1). We deduce the proposition:

PROPOSITION 4.3. *$A^n \otimes A^n$ has a basis as an A-module consisting of the elements*

$$e \otimes Q_{i_1} \dots Q_{i_s} \mathscr{P}^{r_1 \dots r_l} + (-1)^{R+s(n+1)} Q_{i_1} \dots Q_{i_s} \mathscr{P}^{r_1 \dots r_l} \otimes e.$$

Using Lemma 4.2, one proves the proposition in the same way as Proposition 2.3, but by a double induction, first on dimension and then on degree of filtration.

The homomorphism μ of § 2 maps SP^2A^n isomorphically onto the submodule S^2 of $A^n \otimes A^n$ consisting of symmetric tensors. Those elements in (4.3) with $R+s \equiv n \pmod 2$ are symmetric, the remainder

being skew-symmetric; so that the former provide a basis for S^2. Applying ν, which is an isomorphism on S^2, we obtain the corollary:

COROLLARY 4.4. *The module SP^2A^n has as a basis the elements* $e.(Q_{i_1}...Q_{i_s}\mathscr{P}^{r_1...r_l})$ *with* $R+s \equiv n \pmod 2$.

Let $M_2(Z_p, n; Z_p)$ be the submodule of $\bar{H}^*(Z_p, n; Z_p)$ consisting of all elements which can be written in the form $\alpha b.\alpha' b$, with $\alpha, \alpha' \in A$. Then from the results of Cartan mentioned above we have the corollary:

COROLLARY 4.5. *The $(p+1)n$-truncation of $M_2(Z_p, n; Z_p)$ is a truncated free A-module with basis $b.Q_{i_1}...Q_{i_s}\mathscr{P}^{r_1...r_l}b$, where*

$$s + \sum r_i \equiv n \pmod 2.$$

5. The cases $\pi = Z_{p^f}$ and $\pi = Z$

We refer to (**2**) for the definition of the higher-order Bockstein cohomology operations $\beta_m = \beta(p^m)$ $(m > 1)$; recall that these are additive operations raising dimension by one, β_m: $\operatorname{Ker}\beta_{m-1} \to \operatorname{Coker}\beta_{m-1}$, where $\beta_1 = Q_0 \in A$ is the ordinary Bockstein.

Let $\pi = Z_{p^f}$ or Z, and let $b \in H^n(\pi, n; Z_p)$ be the mod p reduction of the basic class. Then according to (**2**), $\beta_m b = 0$ for all m if $\pi = Z$, and for all $m < f$ if $\pi = Z_{p^f}$, in which case $\beta_f b$ is a generator of $H^{n+1}(\pi, n; Z_p) \approx Z_p$.

5.1. *Notation.* We denote by $A(\pi, n)$ the pn-truncation of the submodule of $\bar{H}^*(\pi, n; Z_p)$ which is generated over A by b if $\pi = Z$ or Z_p, and by b and $\beta_f b$ if $\pi = Z_{p^f}$ with $f > 1$.

If we set $A'^n = (A/AQ_0)^n$ and $H^n = A'^n + A'^{n+1}$, then according to (**2**), $A(\pi, n)$ is isomorphic to the pn-truncation of A'^n if $\pi = Z$, of H^n if $\pi = Z_{p^f}$ with $f > 1$.

Let X be a topological space, and let $c \in H^n(X, Z_p)$ be such that $\beta_f c$ is defined, and c and $\beta_f c$ together generate a submodule of $\bar{H}^*(X, Z_p)$ whose pn-truncation M is isomorphic to $A(Z_{p^f}, n)$. Then we shall say that M is isomorphic to $A(Z_{p^f}, n)$ with *generator* c.

Adem and Cartan have given a set of elements D, forming a Z_p-basis for A, with the following properties: $D \supset C \supset B$, where B consists of all allowable iterated Steenrod operations St^I [cf. (**2**)] which neither begin nor end with Q_0; C consists of the elements e, α, and $Q_0\alpha$, where $\alpha \in B$ and $\alpha \neq e$; and D consists of the elements e, Q_0, α, $Q_0\alpha$, αQ_0, and $Q_0\alpha Q_0$ (all of which are non-zero and distinct). The set C therefore furnishes a Z_p-basis for A'^n. With a view to applications we denote the equivalence classes of e in the summands A'^n and A'^{n+1} of H^n by b and β respectively.

Let θ_i denote any element γb or $\gamma\beta$ with $\gamma \in C$. Then $\otimes^m H^n$ has a Z_p-basis consisting of all elements $\theta_1 \otimes ... \otimes \theta_m$, while $SP^m H^n$ has a Z_p-basis consisting of the elements $\theta_1 ... \theta_m$ in which no θ_i of odd dimension is repeated. We have the lemma:

LEMMA 5.2. *$\otimes^m H^n$ has as a module the set of generators $\theta_1 \otimes ... \otimes \theta_m$, where $\theta_1 = b$ or β, and the first θ_i different from b or β is of the form $\theta_\nu = \gamma_\nu b$ or $\gamma_\nu \beta$ with $\gamma_\nu \in B$. The only relations are $Q_0(\theta_1 \otimes ... \otimes \theta_m) = 0$ provided that each θ_i is equal to either b or β.*

Proof. Let N be the submodule of $\otimes^m H^n$ spanned by the elements of the lemma. We shall prove that every monomial $\theta_1 \otimes ... \otimes \theta_m$ is contained in N by induction on $\dim \theta_1$. Suppose that $\theta_1 = \gamma_1 b$ and that the first θ_i different from b or β, other than θ_1, is $\theta_\lambda = \gamma_\lambda b$. Then, if $\gamma_\lambda \in B$, by induction we have

$$\gamma_1(b \otimes \theta_2 \otimes ... \otimes \theta_m) = \theta_1 \otimes ... \otimes \theta_m + \text{elements in } N;$$

while, if $\theta_\lambda = Q_0 \gamma'_\lambda b$, with $\gamma'_\lambda \in B$, then

$$\gamma_1 Q_0(b \otimes \theta_2 \otimes ... \otimes \gamma'_\lambda b \otimes ... \otimes \theta_m) = \pm\theta_1 \otimes ... \otimes \theta_m + \text{elements in } N.$$

In either case this shows that $\theta_1 \otimes ... \otimes \theta_m \in N$. The proof is the same if θ_1 or θ_λ has the form $\gamma\beta$.

Next suppose that there exists a relation

$$\textstyle\sum_{j,l} \alpha_{j,l} \mu_l = 0,$$

where $\alpha_{jl} \in A$ and the μ_l are distinct generators as in Lemma 5.2. The sums involving those μ_l which begin with b and those which begin with β must separately be zero; we consider the first case

$$\mu_l = b \otimes \theta_2^l \otimes ... \otimes \theta_m^l,$$

the second being similar. Rewrite the relation as

$$\textstyle\sum_{j,l} \alpha_{jl} \mu_l + \sum_{j,k} \alpha'_{jk} Q_0 \mu_k = 0, \tag{5.3}$$

where the α_{jl} for fixed l (or the α'_{jk} for fixed k) are either zero or scalar multiples of distinct elements in C. If, for some k, all $\theta_i^k = b$ or β, then $Q_0(\mu_k) = 0$ is a known relation, so we may assume that no such μ_k occur in the second summation.

Let the set C, which provides a Z_p-basis for A'^n, be ordered in a manner compatible with dimension. We may extend this to an ordering of a basis for H^n by requiring, for example, that γb precede $\gamma'\beta$, for all $\gamma, \gamma' \in C$. Finally we order the Z_p-basis elements of $\otimes^m H^n$ lexico-

graphically. If we expand (5.3) in terms of this basis, then the sum of terms of highest order must be zero; but this sum reduces either to

$$(\alpha_{gh}\, b) \otimes \theta_2^h \otimes \ldots \otimes \theta_m^h, \tag{5.4}$$

where α_{gh} is the element of highest order among the coefficients α_{jl}, and μ_h is that element μ_l of highest order which appears in (5.3) with coefficient α_{gh}; or to

$$\pm(\alpha'_{gh}\, b) \otimes \theta_2^h \otimes \ldots \otimes Q_0\, \theta_\nu^h \otimes \ldots \otimes \theta_m^h, \tag{5.5}$$

where α'_{gh} is the element of highest order among the coefficients α'_{jk}, and μ_h is that element μ_k, appearing in (5.3) with coefficient α'_{gh}, such that $b \otimes \theta_2^h \otimes \ldots \otimes Q_0\, \theta_\nu^h \otimes \ldots \otimes \theta_m^h$ has highest order. Here θ_ν^h is as before the first term in μ_h which is not of the form b or β. The elements in (5.4) and (5.5) cannot have equal order since the earliest term in (5.5) (other than the first) which is different from b or β begins with Q_0, while this is not true of (5.4). Hence either $\alpha_{gh} = 0$ or $\alpha'_{gh} = 0$. Proceeding in this fashion, one proves that all coefficients are zero by induction on the order of α_{jl} or α'_{jk}. This completes the proof of the lemma.

From Lemma 5.2 it follows that $\otimes^m H^n$ is a direct sum $S+T$ of submodules, where S has generators $\theta_1 \otimes \ldots \otimes \theta_m$ such that each θ_i is equal to either b or β, and T is A-free. Now $SP^m H^n$ is also a direct sum $S'+T'$ of submodules, where S' is generated over A by $\sigma = b^m$ and $\sigma' = b^{m-1}.\beta$ if n is even, and by $\sigma = b.\beta^{m-1}$ and $\sigma' = \beta^m$ if n is odd, and T' has as a Z_p-basis all of the Z_p-basis elements $\theta_1 \ldots \theta_m$ of $SP^m H^n$ except those of the form $b^{m-1}.\theta$ if n is even, or $\beta^{m-1}.\theta$ if n is odd. Consider the homomorphisms $\mu\colon SP^m H^n \to \otimes^m H^n$ and $\nu\colon \otimes^m H^n \to SP^m H^n$ of § 2, $\nu\mu$ being multiplication by $m!$. Since $\mu(S') \subset S$ and $\nu(S) \subset S'$ if $m < p$, then μ and ν determine $SP^m H^n/S' \approx T'$ as a direct summand of $\otimes^m H^n/S \approx T$. Since T is A-free, so is T'. One shows easily, using μ and ν, that there are no relations in S' other than $Q_0(\sigma) = 0$ and $Q_0(\sigma') = 0$.

Cartan's results (**2**) imply that there is an epimorphism (of Z_p-algebras and A-modules) $\zeta\colon SH^n \to H^*(Z_{p^f}, n; Z_p)$ which induces an isomorphism of pn-truncations, $b \in A'^n$ corresponding to the basic class b, and $\beta \in A'^{n+1}$ to $\beta_f b$. Since the operations β_f are derivations, $\beta_f(b^m) = m b^{m-1}.\beta_f b$ if n is even, and $\beta_f(b.(\beta_f b)^{m-1}) = (\beta_f b)^m$ if n is odd. Hence the image of S' under ζ is isomorphic to $A(Z_{p^f}, mn)$ with generator b^m if n is even, and to $A(Z_{p^f}, mn+m-1)$ with generator

$b.(\beta_f b)^{m-1}$ if n is odd (all modules truncated in dimension pn). Combining the above results for $m = 1,\ldots, p-1$, we have the proposition:

PROPOSITION 5.6. *The pn-truncation of $\bar{H}^*(Z_{p^f}, n; Z_p)$ is the direct sum of the truncations of a free A-module and A-modules G_m ($m = 1,\ldots,p-1$), where*

(i) *if n is even, $G_m \approx A(Z_{p^f}, mn)$ with generator b^m;*
(ii) *if n is odd, $G_m \approx A(Z_{p^f}, mn+m-1)$ with generator $b.(\beta_f b)^{m-1}$.*

The pn-truncations of SA'^n and $\bar{H}^*(Z, n; Z_p)$ are also isomorphic, and reasoning similar to the above yields the proposition:

PROPOSITION 5.7. *The pn-truncation of $\bar{H}^*(Z, n; Z_p)$ is the direct sum of the truncations of a free A-module and the following:*

(i) *if n is even, modules G_m ($m = 1,\ldots,p-1$), where $G_m \approx A(Z, mn)$ with generator b^m;*
(ii) *if n is odd, a module $G_1 \approx A(Z, n)$ with generator b.*

In order to compute generators for SP^2H^n we must change to Milnor's basis for A. We have the lemma:

LEMMA 5.8. *The elements $b \otimes b$ and $b \otimes Q_{i_1}\ldots Q_{i_s}\mathscr{P}^{r_1\ldots r_t}b$, with $r_t \neq 0$ and $1 \leqslant i_1 < \ldots < i_s < t$, form a set of generators for the A-module $A'^n \otimes A'^n$, with the single relation $Q_0(b \otimes b) = 0$.*

Proof. We first show that the submodule N spanned by the given elements is equal to $A'^n \otimes A'^n$. For convenience we allow the sequence $r_1,\ldots, r_t$ to take on the value $0, 0,\ldots,$ corresponding to the element $b \otimes b$. We need only show that $b \otimes \alpha b \in N$ for all $\alpha \in A$ since these elements generate $A'^n \otimes A'^n$. If we can show that $\rho = b \otimes Q_{j_1}\ldots Q_{j_m}\mathscr{P}^{r_1\ldots r_t}b$ is in N provided that $1 \leqslant j_1 < \ldots < j_m$, then

$$Q_0(\rho) = \pm b \otimes Q_0\, Q_{j_1}\ldots Q_{j_m}\mathscr{P}^{r_1\ldots r_t}b$$

is also in N. We prove that $\rho \in N$ by induction on j_m. We may assume that $j_m \geqslant t$ since otherwise $\rho \in N$ by definition. If $j_m = t$, then

$$Q_0(b \otimes Q_{j_1}\ldots Q_{j_{m-1}}\mathscr{P}^{r_1,\ldots,r_t+1}b) = \pm b \otimes Q_{j_1}\ldots Q_{j_{m-1}}\, Q_1\mathscr{P}^{r_1-1,r_2,\ldots,r_t+1}b \pm$$
$$\pm b \otimes Q_{j_1}\ldots Q_{j_{m-1}}\, Q_2\mathscr{P}^{r_1,r_2-1,r_3,\ldots,r_t+1}b \pm \ldots \pm b \otimes Q_{j_1}\ldots Q_{j_m}\mathscr{P}^{r_1,\ldots,r_t}b.$$

Permuting the Q's, we see that each term on the right except the last is either zero or a generator. Hence $\rho \in N$. The proof is similar for $j_m > t$. The relation $Q_0(b \otimes b) = 0$ is obvious.

According to Lemma 5.2, $A'^n \otimes A'^n$ also has generators $b \otimes \gamma b$ ($\gamma \in B$),

with the single relation $Q_0(b \otimes b) = 0$. Now Lemma 8 of (4) implies that there exists a dimension-preserving 1–1 correspondence between the elements of B of positive dimension and the elements

$$Q_{i_1} \ldots Q_{i_s} \mathscr{P}^{r_1 \ldots r_t} \in A$$

with $r_t \neq 0$ and $1 \leqslant i_1 < \ldots < i_s < t$, and hence between the generators of Lemma 5.8 and those of Lemma 5.2 with $m = 2$. Since A, and hence $A'^n \otimes A'^n$, is finite in each dimension, this implies that there can be no further relations among the generators of Lemma 5.8.

From the lemma one easily derives the following corollary, which we shall not use:

COROLLARY 5.9. *A/AQ_0 has a basis as a vectorspace consisting of the elements e and $Q_{i_1} \ldots Q_{i_s} \mathscr{P}^{r_1 \ldots r_t}$ with $r_t \neq 0$ and $0 \leqslant i_1 < \ldots < i_s < t$.*

Henceforth λ_0 is to denote any element of A of the form $Q_{i_1} \ldots Q_{i_s} \mathscr{P}^{r_1 \ldots r_t}$ with $r_t \neq 0$ and $1 \leqslant i_1 < \ldots < i_s < t$, and λ is to denote either e or λ_0. We shall abbreviate $R = \sum r_i$ as before. The filtration in A induces a filtration in $A' = A/AQ_0$, and hence in A'^n and $A'^r \otimes A'^s$. The element λ, considered as lying in either A or A'^n, has filtration precisely $R+s$ (i.e. does not have filtration $R+s-1$). We need the lemma:

LEMMA 5.10. *For n even, the following modules have the indicated generators and relations:*

(1)	$A'^n \otimes A'^n$: $b \otimes \lambda b + (-1)^{R+s} \lambda b \otimes b$,	*relation*	$Q_0(b \otimes b) = 0$,
(2)	$A'^n \otimes A'^{n+1}$: $b \otimes \lambda \beta + (-1)^{R+s} \lambda b \otimes \beta$,	,,	$Q_0(b \otimes \beta) = 0$,
(3)	$A'^{n+1} \otimes A'^n$: $\beta \otimes \lambda b + (-1)^{R} \lambda \beta \otimes b$,	,,	$Q_0(\beta \otimes b) = 0$,
(4)	$A'^{n+1} \otimes A'^{n+1}$: $\beta \otimes \lambda \beta + (-1)^{R} \lambda \beta \otimes \beta$,	,,	$Q_0(\beta \otimes \beta) = 0$.

For n odd, the signs $(-1)^R$ and $(-1)^{R+s}$ should be interchanged. Each of the above elements becomes reducible modulo elements of filtration less than $R+s$ if we change the sign of the second term.

Proof. According to Lemma 4.2, the element $e \otimes \lambda - (-1)^{R+s} \lambda \otimes e$ in $A^n \otimes A^n$ is reducible modulo elements of filtration less than $R+s$; hence the same is true of the element

$$b \otimes \lambda b - (-1)^{R+s} \lambda b \otimes b \in A'^n \otimes A'^n.$$

Part (1) is now proved in the same manner as Proposition 4.3, using Lemma 5.9. The proofs of the remaining parts are similar.

We must find a set of generators for $H^n \otimes H^n$ which is the direct

sum of the modules (1)–(4) in Lemma 5.10, consisting of symmetric and skew-symmetric tensors. The generators of (1) and (4) are already such; we transform those of (2) and (3) by taking sum and difference to obtain, for n even,

(i) $(b\otimes\lambda\beta+(-1)^R\lambda\beta\otimes b)+(\beta\otimes\lambda b+(-1)^{R+s}\lambda b\otimes\beta)$,

(ii) $(b\otimes\lambda\beta-(-1)^R\lambda\beta\otimes b)-(\beta\otimes\lambda b-(-1)^{R+s}\lambda b\otimes\beta)$.

According to Lemma 5.10, these elements with the signs of the second and fourth terms changed, call them (i)′ and (ii)′ respectively, are reducible modulo terms of filtration less than $R+s$. Adding (ii)′ to (i) and (i)′ to (ii) yields the following minimal set of generators for $A'^n\otimes A'^{n+1}+A'^{n+1}\otimes A'^n$, n even:

$$b\otimes\lambda\beta\pm(-1)^R\lambda\beta\otimes b.$$

For n odd, the sign $(-1)^R$ should be replaced by $(-1)^{R+s}$.

We again denote by $M_2(\pi, n;\, Z_p)$, $\pi = Z$ or Z_{p^f}, the submodule of $\bar{H}^*(\pi, n;\, Z_p)$ consisting of all elements which can be written in the form $\theta_1.\theta_2$, with $\theta_i = \alpha b$ or $\alpha\beta_f b$, $\alpha\in A$. Then Lemma 5.10 and the remarks following it yield, in a manner similar to the Corollaries 4.4 and 4.5, the propositions:

Proposition 5.11. *The $(p+1)n$-truncation of $M_2(Z_{p^f}, n;\, Z_p)$ $(f>1)$ is the direct sum of the truncations of an A-module G_2 as in Proposition* 5.6 *and a free A-module with the following basis, where $\lambda_0\in A$ has the form $Q_{i_1}\ldots Q_{i_s}\mathscr{P}^{r_1\ldots r_t}$ with $r_t\neq 0$ and $1\leqslant i_1<\ldots<i_s<t$:*

(i) $b.\lambda_0 b$, $\qquad s+\sum r_i\equiv n \pmod 2$,

(ii) $\beta_f b.\lambda_0\beta_f b$, $\qquad s+\sum r_i\equiv n+1 \pmod 2$,

(iii) $b.\lambda_0\beta_f b$.

If $\pi = Z$, then the calculations reduce to those of the first case in Lemma 5.10 and Proposition 5.11:

Proposition 5.12. *The $(p+1)n$-truncation of $M_2(Z, n;\, Z_p)$ is the direct sum of the truncations of a free A-module with basis $b.\lambda_0 b$, where λ_0 is as in* (5.11) *with $s+\sum r_i\equiv n$* (mod 2), *and* (*provided that n is even*) *a module G_2 as in* (5.7).

The above results are easily generalized to the cohomology

$$H^*(\textstyle\bigtimes_i K(\pi_i, n_i), Z_p)$$

of a product of Eilenberg–MacLane spaces provided that each group

π_i is finitely-generated. Since $K(\pi_1+\pi_2, n)$ has the homotopy type of $K(\pi_1, n)\times K(\pi_2, n)$, we need consider only the case in which each π_i is cyclic of infinite or prime-power order, and we may omit those terms whose orders are relatively prime to p. Let

$$Y = \bigtimes_I K(Z, l_i) \bigtimes_J K(Z_{p^{f_j}}, m_j) \bigtimes_K K(Z_p, n_k),$$

where $(l_i)_{i\in I}$, $(m_j)_{j\in J}$, and $(n_k)_{k\in K}$ are finite sequences and $1 < f_1 \leqslant f_2 \leqslant \dots$. Then

$$H^*(Y, Z_p) = \otimes_I H^*(Z, l_i;\ Z_p) \otimes_J H^*(Z_{p^{f_j}}, m_j;\ Z_p) \otimes_K H^*(Z_p, n_k;\ Z_p). \tag{5.13}$$

Let b_μ be the basic class in the μth term ($\mu \in I$, J, or K). Let $R \subset I$, S, $T \subset J$ be subsets such that m_j is even for $j \in S$, odd for $j \in T$; and let $(\sigma_r)_R$, $(\tau_s)_S$, and $(\tau_t)_T$ be sequences of positive integers such that $\sigma_r = 1$ if l_r is odd. Finally, let $\omega = \min_{I,J,K} l_i, m_j, n_k$; $\nu = \min_{S,T} j$; and let d be the number of elements in $S \cup T$. Then in (5.13) we obtain the following summands:

(1) for each sequence $(\sigma_r)_R$, a summand $A(Z, h)$, $h = \sum_R \sigma_r l_r$, generated by $\otimes_R b_r^{\sigma_r}$;

(2) for each triple $(\sigma_r)_R$, $(\tau_s)_S$, $(\tau_t)_T$, $(d-1)!/\gamma!(d-\gamma-1)!$ summands $A(Z_{p^f}, h+\gamma)$ for $\gamma = 0, 1, \dots, d-1$, where $f = f_\nu$ and

$$h = \textstyle\sum_R \sigma_r l_r + \sum_S \tau_s m_s + \sum_T \tau_t m_t + \tau_t - 1;$$

generators are $\otimes_R b_r^{\sigma_r} \otimes_S x_s \otimes_T y_t$, where $x_s = b_s^{\tau_s}$ and $y_t = b_t.(\beta_{f_t} b_t)^{\tau_t-1}$, with the exception of a total of γ terms which have the form

$$x_s = b_s^{\tau_s-1}.\beta_{f_s} b_s \quad \text{or} \quad y_t = (\beta_{f_t} b_t)^{\tau_t}.$$

The generator with the lowest index (x_ν or y_ν) is not to be among these γ terms. Then we have the proposition:

PROPOSITION 5.14. *The $p\omega$-truncation of $\bar{H}^*(Y, Z_p)$ is the direct sum of the truncations of a free A-module and the A-modules listed in* (1) *and* (2) *above.*

A basis for the free submodule in dimensions less than 3ω can be obtained as iterated tensor products by using Propositions 4.5, 5.11, and 5.12.

6. An application

If $\mathscr{C}$ is a class of Abelian groups (6), then a map will be called an *r-homotopy equivalence* mod $\mathscr{C}$ if it induces isomorphisms mod $\mathscr{C}$ of the homotopy groups in dimensions not exceeding r. $\mathscr{C}_p$ will denote the class of finite Abelian groups with order relatively prime to p.

Let n and q be positive integers, and let X be an $(n-1)$-connected CW complex such that

(i) $\pi_i(X)$ is finitely-generated for $i \leqslant q$,

(ii) cup products in $\bar{H}^*(X, R)$ are zero in dimensions not exceeding q with coefficients in any commutative ring R (e.g. if X is a suspension, then this condition holds for any q);

(iii) the $(q+1)$-truncation of $\bar{H}^*(X, Z_p)$ is isomorphic (as a module over the Steenrod algebra A) to the truncation of a direct sum

$$\sum_{i \in I} A(C_i, n_i) \quad (C_i = Z, Z_p \text{ or } Z_{p^{f_i}}),$$

the isomorphism preserving the operation of the higher-order Bocksteins on the generators e_i of the terms $A(C_i, n_i)$.

Let $\kappa = \min(q, 3n-2)$, and let $Y = \bigtimes_I K(C_i, n_i)$. Then there is a $(\kappa-1)$-homotopy equivalence mod $\mathscr{C}_p$ from X to the space obtained from Y by killing the product terms in the cohomology. More precisely, Proposition 5.14 furnishes an isomorphism of $\kappa+1$-truncations

$$\bar{H}^*(Y, Z_p) = \textstyle\sum_I A(C_i, n_i) + \sum_J A(D_j, m_j),$$

where $(m_j)_J$ is some finite sequence, $D_j = Z$, Z_p, or $Z_{p^{f_j}}$, and the following hold:

(1) $A(C_i, n_i)$ is generated by the class $b_i \in \bar{H}^*(C_i, n_i; Z_p) \subset \bar{H}^*(Y, Z_p)$ which is the mod p reduction of the basic class $b'_i \in H^{n_i}(C_i, n_i; C_i)$.

(2) The generator d_j of $A(D_j, m_j)$ is a (non-trivial) cup or tensor product, and is the mod p reduction of a class $d'_j \in \bar{H}^*(Y, D_j)$ which is also a cup or tensor product.

Let $W = \bigtimes_J K(D_j, m_j)$, and let $g\colon Y \to W$ be a map such that the composition $Y \to W \to K(D_j, m_j)$ realizes d'_j for each $j \in J$. Then g induces a fibre space $p\colon P \to Y$ with fibre ΩW, the space of loops on W.

According to (iii), the generator e_i of the summand $A(C_i, n_i)$ of $\bar{H}^*(X, Z_p)$ is the mod p reduction of a class $e'_i \in \bar{H}^*(X, C_i)$. Let $F_0\colon X \to Y$ be a map such that $F_0^*(b'_i) = e'_i$ for all $i \in I$. Since $F_0^* d'_j$ is a cup product, according to (ii) it is zero, and therefore F_0 lifts to a map $F\colon X \to P$ such that $pF = F_0$. The cohomology sequence of the fibring is exact in dimensions not exceeding κ, and it follows that F induces isomorphisms of the mod p cohomology in these dimensions. Then from Theorem 3 and Proposition 2 [p. 276] of (**6**) we infer the proposition:

Proposition 6.1. *The map $F\colon X \to P$ is a $\kappa-1$-homotopy equivalence* mod $\mathscr{C}_p$.

The elements d'_j form a set of cyclic Postnikov invariants for P in the sense of (1). The map F induces a map of a Postnikov system for X into that for P; and it is easily seen that a set of cyclic invariants χ_j $(j \in J)$, and χ'_k can be chosen for X in dimensions less than κ such that the following hold:

(a) If $D_j = Z_p$ or Z_{p^f}, then d'_j maps into χ_j, which has the same order.

(b) If $D_j = Z$, then d'_j maps into a multiple $q_j \chi_j$, where q_j is relatively prime to p.

(c) The cyclic invariants χ'_k correspond to finite summands of $\pi_*(X)$ which have order relatively prime to p.

6.2. *Example*. Let $X = S^m K(\pi, n)$, the m-fold suspension of an Eilenberg–MacLane space, and let $q = m+pn-1$. Then according to the above discussion, the mod $\mathscr{C}_p$ Postnikov invariants of X are primary operations in dimensions less than κ, where

$$\kappa = \min(m+pn-1,\ 3m+3n-2),$$

and we can write them down explicitly if we compute the number of basis elements, in each dimension less than pn, for the truncated free A-submodule of $\bar{H}^*(\pi, n; Z_p)$. In dimensions less than $3n$ this information can be obtained from Propositions 4.5, 5.11, and 5.12, using a cyclic decomposition for π.

In particular we can conclude that there is a $(\kappa-2)$-homotopy equivalence mod $\mathscr{C}_p$ from $\Omega S^m K(\pi, n)$ to a product of Eilenberg–MacLane spaces since the Postnikov invariants of a loop space are the suspensions of those of the original space, and the suspension of a product is zero. Also, if $m > (p-2)n$, then there is an $(m+pn-2)$-homotopy equivalence mod $\mathscr{C}_p$ from $S^m K(\pi, n)$ to a product of Eilenberg–MacLane spaces; for there can be no products within this dimension range, and therefore all Postnikov invariants are zero.

REFERENCES

1. W. D. Barcus, 'The stable suspension of an Eilenberg-MacLane space', *Trans. American Math. Soc.* 96 (1960) 101–14.
2. H. Cartan, 'Sur les groupes d'Eilenberg–MacLane (II)', *Proc. Nat. Acad. Sci. U.S.A.* 40 (1954) 704–7.
3. —— *Seminaire Henri Cartan (École Normale Supérieure)* (Paris, 1958–9).
4. J. W. Milnor, 'The Steenrod algebra and its dual', *Annals of Math.* 67 (1958), 150–71.
5. —— and J. C. Moore, *On the structure of Hopf algebras* (to appear).
6. J.-P. Serre, 'Groupes d'homotopie et classes de groupes Abéliens', *Annals of Math.* 58 (1953), 258–94.

A CONVEXITY PROPERTY OF POSITIVE MATRICES

By J. F. C. KINGMAN (*Cambridge*)

[Received 16 February 1961]

LET $A(\theta)$ be an $N \times N$ matrix whose elements $a_{ij}(\theta)$ are non-negative functions of the real variable θ. By Frobenius' theorem, $A(\theta)$ has a non-negative eigenvalue $\lambda(\theta)$ with the property that no eigenvalue of $A(\theta)$ exceeds it in modulus. In connexion with a study of Markov chains, Miller (2) has shown, using complex-variable methods, that, if the $a_{ij}(\theta)$ are the Laplace transforms of positive functions, then $\lambda(\theta)$ is a convex function. The purpose of this note is to give a direct proof of a more general result, namely that, if $\log a_{ij}(\theta)$ is convex, then so is $\log \lambda(\theta)$.

In what follows, a 'convex' function will mean a function convex in some fixed interval I: $\theta_1 < \theta < \theta_2$.

We say that a positive function $f(\theta)$ is *superconvex* if $\log f(\theta)$ is convex. Clearly a positive function $f(\theta)$ is superconvex if and only if, for all θ, ϕ in I, and all $\alpha, \beta \geqslant 0$ with $\alpha+\beta = 1$,

$$f(\alpha\theta+\beta\phi) \leqslant f(\theta)^{\alpha} f(\phi)^{\beta}. \tag{1}$$

Denote by $\mathfrak{S}$ the class of all superconvex functions, together with the function identically zero in I. Then $\mathfrak{S}$ is the class of functions satisfying (1) since, if $f(\theta) = 0$, then $f(\alpha\theta+\beta\phi) = 0$ for all α, β, ϕ; so that $f \equiv 0$ in I.

LEMMA. *$\mathfrak{S}$ is closed under addition, multiplication, and raising to any positive power. If, for each n, f_n belongs to $\mathfrak{S}$, then so does $f = \limsup f_n$ as n tends to infinity.*

Proof. Let f, g satisfy (1), and let $h = f+g$. Then

$$\begin{aligned} h(\alpha\theta+\beta\phi) &= f(\alpha\theta+\beta\phi)+g(\alpha\theta+\beta\phi) \\ &\leqslant f(\theta)^{\alpha} f(\phi)^{\beta}+g(\theta)^{\alpha} g(\phi)^{\beta} \\ &\leqslant \{f(\theta)+g(\theta)\}^{\alpha}\{f(\phi)+g(\phi)\}^{\beta}, \text{ by Hölder's inequality [(1) 22]}, \\ &= h(\theta)^{\alpha} h(\phi)^{\beta}. \end{aligned}$$

Hence $\mathfrak{S}$ is closed under addition. It is also closed under multiplication and raising to positive powers since any linear combination of convex functions with positive coefficients is convex.

If the f_n satisfy (1), and if $f = \limsup f_n$, then

$$\begin{aligned} f(\alpha\theta+\beta\phi) &\leqslant \limsup f_n(\theta)^\alpha f_n(\phi)^\beta \\ &\leqslant \limsup f_n(\theta)^\alpha . \limsup f_n(\phi)^\beta \\ &= f(\theta)^\alpha f(\phi)^\beta. \end{aligned}$$

Hence f belongs to $\mathfrak{S}$, and the lemma is proved.

THEOREM. *Let $A(\theta)$ be an $N \times N$ matrix such that $a_{ij}(\theta)$ belongs to $\mathfrak{S}$. Then its largest eigenvalue $\lambda(\theta)$ belongs to $\mathfrak{S}$.*

Proof. Let $f_n(\theta) = \{\text{trace}\, A^n(\theta)\}^{1/n}$.

From the lemma it follows that f_n belongs to $\mathfrak{S}$. But, if $\lambda_1(\theta)$, $\lambda_2(\theta)$,... are the eigenvalues of $A(\theta)$, then

$$f_n(\theta) = \{\textstyle\sum \lambda_j^n(\theta)\}^{1/n}.$$

Thus $\lambda(\theta) = \limsup f_n(\theta)$, and hence $\lambda(\theta)$ belongs to $\mathfrak{S}$.

COROLLARY 1. *Since any superconvex function is convex, $\lambda(\theta)$ is convex.*

COROLLARY 2. *The theorem holds if*

$$a_{ij}(\theta) = \int_{-\infty}^{\infty} e^{\theta x}\, dF_{ij}(x),$$

where $F_{ij}(x)$ is increasing, and each a_{ij} converges in I.

This follows from the inequality

$$\int e^{(\alpha\theta+\beta\phi)x}\, dF(x) \leqslant \Big\{\int e^{\theta x}\, dF(x)\Big\}^\alpha \Big\{\int e^{\phi x}\, dF(x)\Big\}^\beta,$$

which is a consequence of Hölder's inequality [cf. (1) 140].

These two corollaries yield Miller's result. It may be noted that the theorem holds for any class of functions satisfying the lemma. However, it is not true that, if the a_{ij} are convex, then so is $\lambda(\theta)$. In fact, the class of convex functions is not closed under multiplication.

REFERENCES

1. G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities* (Cambridge, 1959).
2. H. D. Miller, 'A convexity property in the theory of random variables defined on a finite Markov chain', *Ann. Math. Stat.* (to appear).

SOME IDENTITIES IN COMBINATORIAL ANALYSIS

By B. GORDON (*Los Angeles*)

[Received 14 April 1961]

1. Introduction

A FAMOUS identity due to Jacobi in the theory of elliptic functions says that, if x and z are complex variables with $|x| < 1$ and $z \neq 0$, then

$$\prod_{n=1}^{\infty}(1-x^{2n})(1+x^{2n-1}z)(1+x^{2n-1}z^{-1}) = \sum_{-\infty}^{\infty} x^{n^2}z^n \tag{1}$$

[for a proof cf. (1) 282]. This identity has numerous consequences of interest in number theory and combinatorial analysis, among them the formulae

$$\prod_{n=1}^{\infty}(1-x^n) = \sum_{-\infty}^{\infty}(-1)^n x^{\frac{1}{2}(3n^2+n)}, \tag{2}$$

$$\prod_{n=1}^{\infty}(1-x^n)^3 = \sum_{n=0}^{\infty}(-1)^n(2n+1)x^{\frac{1}{2}(n^2+n)}, \tag{3}$$

$$\prod_{n=1}^{\infty}(1-x^{2n})/(1-x^{2n-1}) = \sum_{n=0}^{\infty} x^{\frac{1}{2}(n^2+n)}, \tag{4}$$

due to Euler, Jacobi, and Gauss respectively [cf. (1) 283–5]. In the present paper I obtain an identity analogous to (1), viz.

$$\prod_{n=1}^{\infty}(1-x^{2n})(1-x^{2n-1}z)(1-x^{2n-1}z^{-1})(1-x^{4n-4}z^2)(1-x^{4n-4}z^{-2})$$
$$= \sum_{-\infty}^{\infty} x^{3n^2-2n}(z^{3n}+z^{-3n}-z^{3n-2}-z^{-3n+2}), \tag{5}$$

valid for $|x| < 1$ and $z \neq 0$. From this it is possible to derive new formulae of the type of (2), (3), (4), a typical example being

$$\prod_{n=1}^{\infty}(1-x^n)^3(1-x^{2n-1})^2 = \sum_{-\infty}^{\infty}(6n+1)x^{\frac{1}{2}(3n^2+n)}.$$

Finally, we obtain some congruence properties analogous to those possessed by the partition function $p(n)$ defined by

$$f(x) = \prod_{n=1}^{\infty}(1-x^n)^{-1} = \sum_{n=0}^{\infty} p(n)x^n.$$

For example, if $$f(x)^2 f(x^2)^2 = \sum_{n=0}^{\infty} c_n x^n,$$

then $c_n \equiv 0 \pmod 7$ whenever $n \equiv 2, 3, 4,$ or $6 \pmod 7$.

2. Proof of (5)

It is convenient to prove first the identity

$$\prod_{n=1}^{\infty}(1-s^n)(1-s^nt)(1-s^{n-1}t^{-1})(1-s^{2n-1}t^2)(1-s^{2n-1}t^{-2})$$
$$= \sum_{-\infty}^{\infty} s^{\frac{1}{2}(3n^2+n)}(t^{3n}-t^{-3n-1}), \quad (6)$$

from which (5) results by putting $s = x^2$, $t = x^{-1}z$, and multiplying through by $1-z^{-2}$. The product $A(s,t)$ on the left of (6) is absolutely convergent for $|s| < 1$, $t \neq 0$, and can be written in the form

$$A(s,t) = \sum_{-\infty}^{\infty} \phi_n(s)t^n, \quad (7)$$

the series converging uniformly for $|s| \leqslant \theta < 1$ and $T^{-1} \leqslant |t| \leqslant T$. Putting

$$B(s,t) = \prod_{n=1}^{\infty}(1-s^nt),$$

we have

$$A(s,t) = (1-t^{-1})B(s,t)B(s,t^{-1})B(s^2,s^{-1}t)B(s^2,s^{-1}t^{-1})B(s,1).$$

Since
$$B(s,t) = (1-st)B(s,st),$$
an easy calculation shows that

$$A(s,t) = s^2t^3A(s,st).$$

Substituting this into (7) we find that

$$\phi_n(s) = s^{n-1}\phi_{n-3}(s), \quad (8)$$

and so all the coefficients $\phi_n(s)$ can be determined once ϕ_{-1}, ϕ_0, and ϕ_1 are known. On the other hand,

$$A(s,t)/(1-t^{-1}) = \sum_{-\infty}^{\infty} \psi_n(s)t^n,$$

where $\phi_n = \psi_n - \psi_{n+1}$. Since $A(s,t)/(1-t^{-1})$ is invariant under the substitution $t \to t^{-1}$, we have $\psi_n = \psi_{-n}$. From the equations

$$\phi_{-1} = \psi_{-1} - \psi_0, \qquad \phi_0 = \psi_0 - \psi_1$$

we thus obtain $\phi_{-1} = -\phi_0$. Similarly, from

$$\phi_1 = \psi_1 - \psi_2, \qquad \phi_{-2} = \psi_{-2} - \psi_{-1}, \qquad \phi_1 = \phi_{-2}$$

we get $\phi_1 = 0$. Hence, using (8), we see that

$$A(s,t) = \phi_0(s)H(s,t),$$

where $H(s,t)$ is the series on the right of (6). To evaluate ϕ_0, set

$$t = \zeta = \exp(\tfrac{1}{3}i\pi),$$

a primitive sixth root of unity. Then a simple calculation yields

$$H(s,\zeta) = (1-\zeta^{-1})\sum_{-\infty}^{\infty}(-1)^n s^{\frac{1}{2}(3n^2+n)} = (1-\zeta^{-1})\prod_{n=1}^{\infty}(1-s^n)$$

by Euler's identity (2). On the other hand

$$A(s,\zeta) = (1-\zeta^{-1})\prod_{n=1}^{\infty}(1-s^n)(1-s^n\zeta)(1-s^n\zeta^{-1})(1-s^{2n-1}\zeta^2)(1-s^{2n-1}\zeta^{-2})$$

from which it follows that

$$\phi_0(s) = \prod_{n=1}^{\infty}(1-s^n\zeta)(1-s^n\zeta^{-1})(1-s^{2n-1}\zeta^2)(1-s^{2n-1}\zeta^{-2}).$$

Now

$$\begin{aligned}\omega(s) &= \prod_{n=1}^{\infty}(1-s^n\zeta)(1-s^{2n-1}\zeta^2)\\ &= \prod_{n=1}^{\infty}(1-s^n\zeta)(1-s^{2n-1}\zeta^2)(1-s^{2n}\zeta^2)/(1-s^{2n}\zeta^2)\\ &= \prod_{n=1}^{\infty}(1-s^n\zeta^2)/(1+s^n\zeta).\end{aligned}$$

If s is real, each factor of this product is on the unit circle (since $\zeta^2 = -\bar{\zeta}$); hence so is the product itself. Thus for real s,

$$\phi_0(s) = \omega(s)\overline{\omega(s)} = 1,$$

and by analytic continuation, $\phi_0(s) = 1$ for all s with $|s| < 1$. This completes the proof of (6) and so also of (5).

3. Applications

In (6) set $s = x^M$, $t = x^{-k}$, where M and k are integers satisfying $0 < 2k < M$. We then obtain

$$\prod_{n=1}^{\infty}(1-x^n) = \sum_{-\infty}^{\infty} x^{M\frac{1}{2}(3n^2+n)}(x^{-3nk}-x^{3nk+k}),$$

where n is 0, $\pm k$, M, $M\pm k$, $M\pm 2k \pmod{2M}$ and where the factor $1-x^n$ is taken twice on the left if n is in two of the permitted residue classes $(\mathrm{mod}\, 2M)$. The right-hand side can be rewritten in the form

$$\sum_{-\infty}^{\infty}(-1)^n x^{kT_n+(M-4k)U_n},$$

where $$T_n = \tfrac{1}{2}(3n^2+n), \qquad U_n = T_{[\frac{1}{2}(n+1)]}.$$

This gives a generalization of Euler's identity (2), to which it reduces when $M = 4$, $k = 1$. If $M = 3$, $k = 1$, we get

$$\prod_{n=1}^{\infty}(1-x^n)(1-x^{6n-5})(1-x^{6n-1}) = 1-2x+x^3+x^6-2x^{10}+\ldots.$$

Putting $\quad f(x) = \prod_{n=1}^{\infty}(1-x^n)^{-1}, \qquad g(x) = \sum_{n=0}^{\infty} x^{\frac{1}{2}(n^2+n)},$

we can write this in the form

$$f(x^2)f(x^3)/f(x)^2f(x^6) = g(x)-3xg(x^9). \tag{9}$$

A formula equivalent to (9) appears several times in Ramanujan's notebooks (**2**). Since

$$g(x) = \prod_{n=1}^{\infty}(1-x^{2n})/(1-x^{2n-1}) = f(x)/f(x^2)^2$$

by Gauss's identity (4), this leads to a functional equation for $f(x)$. Again, if $M = 6$, $k = 1$ we find in the same manner that

$$2f(x^2)f(x^3)f(x^{12})/f(x)f(x^4)f(x^6)^2 = 3\theta(x^9)-\theta(x), \tag{10}$$

where $\theta(x) = \sum_{-\infty}^{\infty} x^{n^2}$. From this one can deduce functional equations for either $f(x)$ or $\theta(x)$. A curious corollary is the fact that the equation $\theta(x) = 3\theta(x^9)$ has no solutions.

If in formula (6) we divide by $1-t^{-1}$ and then let $t \to 1$, we obtain

$$\prod_{n=1}^{\infty}(1-s^n)^3(1-s^{2n-1})^2 = \sum_{-\infty}^{\infty}(6n+1)s^{\frac{1}{2}(3n^2+n)}. \tag{11}$$

This formal procedure is easily justified as in the case of Jacobi's identity (3) [for details cf. (**1**) 285].

In the same way we find by dividing (5) by $(1-z^2)(1-z^{-2})$ and then letting $z \to 1$ that

$$\prod_{n=1}^{\infty}(1-x^{2n})(1-x^{2n-1})^2(1-x^{4n})^2 = \sum_{-\infty}^{\infty}(3n+1)x^{3n^2+2n}. \tag{12}$$

4. Congruences

Let p be a prime and suppose that

$$A(x) = \sum_{n=0}^{\infty} a_n x^n$$

is such that $a_n \equiv 0 \pmod{p}$ whenever $n \equiv n_0 \pmod{p}$. If

$$B(x) = \sum_{n=0}^{\infty} b_n x^n$$

has $b_n \equiv 0 \pmod{p}$ for all $n \not\equiv 0 \pmod{p}$, then

$$A(x)B(x) = \sum_{n=0}^{\infty} c_n x^n$$

clearly satisfies $c_n \equiv 0 \pmod{p}$ for all $n \equiv n_0 \pmod{p}$. Since the pth power of any

$$D(x) = \sum_{n=0}^{\infty} d_n x^n$$

with integer coefficients is a suitable $B(x)$, we can obtain a number of congruences from the identities discussed in § 3. A few examples should serve to illustrate this. Taking $p = 3$, $A(x)$ the function of (10), $n_0 = 2$, and $B(x) = f(x^6)^2/f(x^3)f(x^{12})$, we see that

$$f(x^2)/f(x)f(x^4) = \sum_{n=0}^{\infty} c_n x^n$$

has $c_n \equiv 0 \pmod 3$ whenever $n \equiv 2 \pmod 3$.

If $p = 7$, $A(x)$ is the function of (11), $n_0 = 2$, 3, 4, or 6, and

$$B(x) = \prod_{n=1}^{\infty} (1-x^n)^{-7},$$

we see that

$$f(x)^2 f(x^2)^2 = \sum_{n=0}^{\infty} c_n x^n$$

has $c_n \equiv 0 \pmod 7$ whenever $n \equiv 2$, 3, 4, or 6 (mod 7).

If $p = 5$ and

$$A(x) = \prod_{n=1}^{\infty} (1-x^n) \sum_{\nu=-\infty}^{\infty} (3\nu+1)x^{3\nu^2+2\nu} = \sum_{n=0}^{\infty} a_n x^n,$$

we see easily that $a_n \equiv 0 \pmod 5$ whenever $n \equiv 4 \pmod 5$. Hence, multiplying by

$$B(x) = \prod_{n=1}^{\infty} (1-x^n)^{-5}$$

and using (12), we find that

$$f(x^2)f(x)^2/f(x^4)^2 = \sum_{n=0}^{\infty} c_n x^n$$

has $c_n \equiv 0 \pmod 5$ whenever $n \equiv 4 \pmod 5$.

Finally, if $p = 11$, it is readily seen that

$$\left\{\sum_{-\infty}^{\infty} (3n+1)x^{3n^2+2n}\right\}^2 = \sum_{n=0}^{\infty} a_n x^n$$

has $a_n \equiv 0 \pmod{11}$ whenever $n \equiv 7 \pmod{11}$. Hence, taking

$$B(x) = \prod_{n=1}^{\infty} (1-x^n)^{-11}$$

and using (12), we see that

$$f(x^2)^2 f(x)^7/f(x^4)^4 = \sum_{n=0}^{\infty} c_n x^n$$

has $c_n \equiv 0 \pmod{11}$ whenever $n \equiv 7 \pmod{11}$.

5. Possible generalizations

In view of (1) and (5) it is natural to ask for analogous identities involving products of any number of factors. I discuss this problem in terms of (6), which is equivalent to (5) but easier to handle, and of

$$\prod_{n=1}^{\infty}(1-s^n)(1-s^n t)(1-s^{n-1}t^{-1}) = \sum_{-\infty}^{\infty}(-1)^n s^{\frac{1}{2}(n^2+n)} t^n, \tag{13}$$

the corresponding transformation of (1). We put

$$C(s,t) = (1-t^{-1})B(s,t)B(s,t^{-1}), \qquad D(s,t) = B(s^2, s^{-1}t^2)B(s^2, s^{-1}t^{-2}),$$

and consider the function

$$F(s,t) = \prod_{i,j,k,l} C(s^{\alpha_i}, t^{\alpha_i})C(s^{\beta_j}, -t^{\beta_j})D(s^{\gamma_k}, t^{\gamma_k})D(s^{\delta_l}, -t^{\delta_l})$$

($1 \leqslant i \leqslant I$; $1 \leqslant j \leqslant J$; $1 \leqslant k \leqslant K$, $1 \leqslant l \leqslant L$), where α_i, β_j, γ_k, δ_l are positive integers not necessarily distinct. This function is defined for $|s| < 1$, $t \neq 0$, and satisfies the equation

$$F(s,t) = (-1)^{I+K} s^{\alpha+\beta+\gamma+\delta} t^{\alpha+\beta+2\gamma+2\delta} F(s, st),$$

where $\alpha = \sum \alpha_i$, $\beta = \sum \beta_j$, $\gamma = \sum \gamma_k$, $\delta = \sum \delta_l$. We can write

$$F(s,t) = \sum_{-\infty}^{\infty} h_n t^n,$$

where $h_n = h_n(s)$ depends only on s, and the Laurent series converges for $t \neq 0$, $|s| < 1$. From the functional equation we obtain the recurrence formula

$$h_n = (-1)^{I+K} s^{n-\gamma-\delta} h_{n-\epsilon},$$

where $\epsilon = \alpha+\beta+2\gamma+2\delta$. Thus all h_n are determined by $h_0, \ldots, h_{\epsilon-1}$. From the fact that

$$F(s,t)\prod_{i,j}(1-t^{-\alpha_i})^{-1}(1+t^{-\beta_j})^{-1}$$

is invariant under the substitution $t \to t^{-1}$, we find that

$$h_n = (-1)^I h_{-n-\alpha-\beta},$$

which can be used to reduce the number of independent h_n.

If $\epsilon \leqslant 3$, $\alpha+\beta \leqslant 1$, $\beta+\delta \leqslant 2$, then everything can be reduced to the evaluation of a single h_n, and we obtain identities equivalent to (6) or (13). But otherwise there will be at least two independent h_n to evaluate, and I have not succeeded in finding any case where this can be done in an elementary manner.

REFERENCES

1. G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers* (4th ed., Oxford, 1960).
2. S. Ramanujan, *Notebooks* (2 vols., Bombay, 1957).

EIGENFUNCTION EXPANSIONS ASSOCIATED WITH A COMPLEX DIFFERENTIAL OPERATOR OF THE SECOND ORDER

By J. B. McLEOD (*Oxford*)

[Received 23 June 1961]

1. The theory of eigenfunction expansions associated with the real differential equation

$$\frac{d^2y}{dx^2}+\{\lambda-q(x)\}y = 0 \quad (0 \leqslant x < \infty) \tag{1.1}$$

and real self-adjoint boundary conditions has been thoroughly worked out (**1**). The situation is quite different when $q(x)$ and the boundary conditions are allowed to become complex. In this case the differential equation is no longer self-adjoint, and the problem is a particular case of the spectral expansion associated with a non-self-adjoint linear operator. Research in this field is still extremely fragmentary, and it is not to be expected that we can develop eigenfunction expansions associated with the complex case at all as systematically as Titchmarsh has done with the real case.

There are two types of eigenfunction problem to be considered. The first is the non-singular or Sturm–Liouville case, where $q(x)$ is continuous and the boundary conditions are taken at finite end-points,

$$y(a)\cos\alpha+y'(a)\sin\alpha = 0, \tag{1.2}$$

$$y(b)\cos\beta+y'(b)\sin\beta = 0, \tag{1.3}$$

where α, β are complex constants.

The second type is the singular case, where the range of x becomes infinite and/or $q(x)$ becomes discontinuous at one or other or both of the end-points.

The Sturm–Liouville case is one in which even the non-self-adjoint problem can be solved completely. This has been recognized for some time, notably by authors like Birkhoff and Tamarkin [(**2**)–(**6**)], but these authors consider a differential equation of order n, along with boundary conditions much more general than (1.2), (1.3), and as a consequence they are unable to discuss the detailed nature of the

expansion for the case in which we are interested. Peierls (**7**) has gone into details, but his treatment, he admits, is non-rigorous. As Dolph (**8**) points out, the promised justification of this has never appeared. Schwartz (**9**), in a very important and powerful paper, treats the Sturm–Liouville case (and also certain singular cases), but only as a special case of a long and complicated function-theoretic argument. Keldysh (**10**) has also given a linear-operator approach to the problem.

Altogether, there does seem a case for a direct justification of Peierls's work that does not depend on function-theoretic arguments, and this is particularly so when it appears that, without any great complication, it is possible at the same time to make a contribution to the singular case in which the range of x remains finite but $q(x)$ becomes discontinuous at one or other or both of the end-points. This contribution does not seem to be covered by the existing function-theoretic arguments.

The problem we shall consider is the following. We take the equation

$$\frac{d^2y}{dr^2}+\left\{\lambda-q(r)-\frac{l(l+1)}{r^2}\right\}y=0, \tag{1.4}$$

where $q(r)$ may be complex but is continuous except at $r = 0$, and where $\int\limits_0 r|q(r)|\,dr$ exists. We suppose that l is a positive integer or zero. The reader will readily verify that the analysis is not restricted to those values of l, but this is the case of practical importance. (The equation is the well-known equation that arises when a three-dimensional equation with spherical symmetry is solved by the method of separation of variables.)

The boundary conditions we impose are (1.3), for some $b > 0$, together with the requirement that $y(x)$ be $L^2(0, b)$. This, as the analysis shows, is sufficient to define an eigenvalue problem, except in the case $l = 0$, when we have to impose a further condition of the type (1.2) at $a = 0$. Despite this, the case $l = 0$ is similar enough to the case $l > 0$, so that we can safely restrict ourselves to $l > 0$. The case $l = 0$, with $q(r)$ continuous, is just the Sturm–Liouville case, which therefore comes out as a particular case of the argument.

We shall examine the eigenfunctions associated with this eigenvalue problem. As usual, an eigenfunction is a non-trivial solution of the equation (1.4) which satisfies the boundary conditions. In the self-adjoint case, the set of eigenfunctions would be complete, i.e. any reasonable function could be expanded in a series of them. In the non-self-adjoint case, we shall see that in general this no longer holds,

but that the set of eigenfunctions can be made complete by adding to it certain other functions which, though not eigenfunctions, are related to them. (Their precise form will be found in § 5.) I shall refer to these additional functions as *adjoint* functions.

The problem can be extended to the case in which $r = b$ is also a discontinuity of $q(r)$, of the same type as at $r = 0$. It will not be necessary to discuss in detail this extension, but it will be clear that the same general conclusions hold on the completeness of the set of eigenfunctions and adjoint functions.

I have limited myself to proving completeness, but, at least in certain cases, much more can be proved. For example, in the Sturm–Liouville case, a very straightforward adaptation of (1) [Ch. I] shows that not only is the set of eigenfunctions and adjoint functions complete, but also that, if $f(r)$ is any function of $L(0,b)$, then the eigenfunction expansion of $f(r)$ (an expansion which, of course, includes adjoint functions) converges under Fourier conditions to $f(r)$. This analysis does not seem to extend to the singular cases considered in this paper.

2. If $q(r) \equiv 0$, then (1.4) has solutions $r^{\frac{1}{2}}J_{\pm(l+\frac{1}{2})}(r\sqrt{\lambda})$, of which $r^{\frac{1}{2}}J_{l+\frac{1}{2}}(r\sqrt{\lambda})$ is $L^2(0,b)$. If we then write (1.4) in the form

$$\frac{d^2y}{dr^2}+\left\{\lambda-\frac{l(l+1)}{r^2}\right\}y = q(r)y,$$

we see that it is formally equivalent to the integral equation

$$y(r,\lambda) = r^{\frac{1}{2}}J_{l+\frac{1}{2}}(r\sqrt{\lambda})+\tfrac{1}{2}\pi(-1)^{l+1}\times$$
$$\times\int\limits_0^r \{J_{l+\frac{1}{2}}(r\sqrt{\lambda})J_{-l-\frac{1}{2}}(t\sqrt{\lambda})-J_{l+\frac{1}{2}}(t\sqrt{\lambda})J_{-l-\frac{1}{2}}(r\sqrt{\lambda})\}r^{\frac{1}{2}}t^{\frac{1}{2}}q(t)y(t,\lambda)\,dt. \tag{2.1}$$

Our first objective is to prove that, for $0 \leqslant r \leqslant b$, and all $|\lambda|$ sufficiently large, the solution of (1.4) that is $L^2(0,b)$ is, apart from a multiplicative constant,

$$r^{\frac{1}{2}}J_{l+\frac{1}{2}}(r\sqrt{\lambda})+o(\lambda^{-\frac{1}{2}}e^{|\tau|r}), \tag{2.1 a}$$

where $o(1)$ denotes a term small where $|\lambda|$ is large, uniformly for r in $[0,b]$, and where $\kappa = \sqrt{\lambda} = \sigma+i\tau$. We do this by investigating (2.1).

Let $$\omega(r,\lambda) = \begin{cases} r^{l+1}|\lambda|^{\frac{1}{2}l+\frac{1}{2}} & (r \leqslant |\lambda|^{-\frac{1}{2}}), \\ |\lambda|^{-\frac{1}{2}}e^{|\tau|r} & (r > |\lambda|^{-\frac{1}{2}}). \end{cases}$$

Then $$|r^{\frac{1}{2}}J_{l+\frac{1}{2}}(r\sqrt{\lambda})| < A\omega(r,\lambda)$$

for all r, λ, where A denotes various positive constants. Let

$$\max_{0 \leqslant r \leqslant b} \frac{|y(r,\lambda)|}{\omega(r,\lambda)} = M.$$

Then, if $r \leqslant |\lambda|^{-\frac{1}{2}}$, (2.1) gives

$$M < A + AM \int_0^r \left(\frac{r}{t}\right)^{l+\frac{1}{2}} r^{\frac{1}{2}} t^{\frac{1}{2}} \frac{\omega(t,\lambda)}{\omega(r,\lambda)} |q(t)|\, dt = A + AM\, o(1) \qquad (2.1\,\mathrm{b})$$

since $\int\limits_0 t|q(t)|\,dt$ exists, the $o(1)$ term denoting a quantity which tends to zero as $|\lambda| \to \infty$. Also, if $r > |\lambda|^{-\frac{1}{2}}$,

$$M < A + AM \int_0^{|\lambda|^{-\frac{1}{2}}} t|q(t)|\,dt + AM \int_{|\lambda|^{-\frac{1}{2}}}^{r} r^{\frac{1}{2}} t^{\frac{1}{2}} \frac{\omega(t,\lambda)}{\omega(r,\lambda)} G(r,t,\lambda)|q(t)|\,dt, \qquad (2.2)$$

where $\quad G(r,t,\lambda) = |J_{l+\frac{1}{2}}(r\sqrt{\lambda})J_{-l-\frac{1}{2}}(t\sqrt{\lambda}) - J_{-l-\frac{1}{2}}(r\sqrt{\lambda})J_{l+\frac{1}{2}}(t\sqrt{\lambda})|. \qquad (2.3)$

But, for $|r\sqrt{\lambda}| \geqslant |t\sqrt{\lambda}| \geqslant 1$, we have $r^{\frac{1}{2}}t^{\frac{1}{2}}G(r,t,\lambda) < A|\lambda|^{-\frac{1}{2}}e^{|\tau|(r-t)}$. For, for all z,

$$J_{l+\frac{1}{2}}(z) = \tfrac{1}{2}\{H^{(1)}_{l+\frac{1}{2}}(z) + H^{(2)}_{l+\frac{1}{2}}(z)\},$$

$$J_{-l-\frac{1}{2}}(z) = \tfrac{1}{2}(-1)^l i\{H^{(1)}_{l+\frac{1}{2}}(z) - H^{(2)}_{l+\frac{1}{2}}(z)\},$$

so that

$$G(r,t,\lambda) < A\{|H^{(1)}_{l+\frac{1}{2}}(r\sqrt{\lambda})H^{(2)}_{l+\frac{1}{2}}(t\sqrt{\lambda})| + |H^{(2)}_{l+\frac{1}{2}}(r\sqrt{\lambda})H^{(1)}_{l+\frac{1}{2}}(t\sqrt{\lambda})|\}.$$

The required estimate for $G(r,t,\lambda)$ follows from this by using the asymptotic expressions for $H^{(1)}_\nu(z)$, $H^{(2)}_\nu(z)$.

Substituting this estimate in (2.2), we obtain

$$M < A + AM\, o(1) + AM|\lambda|^{-\frac{1}{2}} \int_{|\lambda|^{-\frac{1}{2}}}^{r} |q(t)|\,dt$$

$$= A + AM\, o(1) + AM|\lambda|^{-\frac{1}{2}} \left\{ \int_{|\lambda|^{-\frac{1}{2}}}^{|\lambda|^{-\frac{1}{4}}} + \int_{|\lambda|^{-\frac{1}{4}}}^{r} \right\} \frac{t|q(t)|}{t}\,dt. \qquad (2.4)$$

The first of the two integrals in the last line is $o(|\lambda|^{\frac{1}{2}})$ since $t^{-1} \leqslant |\lambda|^{\frac{1}{2}}$ in the range of integration and

$$\int_{|\lambda|^{-\frac{1}{2}}}^{|\lambda|^{-\frac{1}{4}}} t|q(t)|\,dt = o(1).$$

The second integral is $O(|\lambda|^{\frac{1}{4}})$, by a similar type of argument. (The second integral will not, of course, appear if $r \leqslant |\lambda|^{-\frac{1}{4}}$.)

It thus follows from (2.1 b) and (2.4) that, for $0 \leqslant r \leqslant b$, $M < A$ if $|\lambda|$ is large enough, i.e. that $|y(r,\lambda)| < A\omega(r,\lambda)$. If we substitute this result

back in the integral in (2.1) and re-estimate this integral on the same lines as has just been done, we emerge with (2.1 a).

Thus any solution of (2.1) satisfies (2.1 a). That there is one (and just one) solution of (2.1) can be proved by the usual iteration process, of which the work above is effectively the first step. Then (2.1) can be differentiated back to show that the solution is a solution of (1.4).

We have thus found a solution of (1.4) that is $L^2(0,b)$. If we denote this solution by $\phi(r,\lambda)$, then any other solution apart from a constant multiple of $\phi(r,\lambda)$ is given by a constant multiple of

$$\psi(r,\lambda) = \phi(r,\lambda)\int^{r} \frac{dt}{\phi^2(t,\lambda)},$$

and knowing now the behaviour of $\phi(r,\lambda)$ near $r = 0$, we can readily verify that $\psi(r,\lambda)$ is not $L^2(0,b)$. The $L^2(0,b)$ solution is therefore (apart from a multiplicative constant) unique.

We remark finally that, since $\lambda^{-\frac{1}{2}l-\frac{1}{4}}r^{\frac{1}{2}}J_{l+\frac{1}{2}}(r\sqrt{\lambda})$ is an integral function of λ, the process of solving (2.1) by iteration shows that $\lambda^{-\frac{1}{2}l-\frac{1}{4}}\phi(r,\lambda)$ is also an integral function of λ.

3. We now consider the solution $\chi(r,\lambda)$ which satisfies (1.4) and the boundary conditions

$$\chi(b,\lambda) = \sin\beta, \qquad \chi'(b,\lambda) = -\cos\beta.$$

As in (**1**) [Ch. I] $\chi(r,\lambda)$ is an integral function of λ.

The Wronskian of ϕ, χ is independent of r and so may be written as $\omega(\lambda)$, and $\lambda^{-\frac{1}{2}l-\frac{1}{4}}\omega(\lambda)$ will be an integral function of λ.

Further, the vanishing of $\omega(\lambda)$ is a necessary and sufficient condition for ϕ, χ to be multiples the one of the other, i.e. for λ to be an eigenvalue.

For large values of $|\lambda|$,

$$\omega(\lambda) = \phi(b,\lambda)\chi'(b,\lambda)-\phi'(b,\lambda)\chi(b,\lambda)$$
$$\sim -\cos\beta\lambda^{-\frac{1}{4}}\left(\frac{2}{\pi}\right)^{\frac{1}{2}}\cos(b\sqrt{\lambda}-\tfrac{1}{2}l\pi-\tfrac{1}{2}\pi)+$$
$$+\sin\beta\lambda^{\frac{1}{4}}\left(\frac{2}{\pi}\right)^{\frac{1}{2}}\sin(b\sqrt{\lambda}-\tfrac{1}{2}l\pi-\tfrac{1}{2}\pi). \qquad (3.1)$$

(The asymptotic behaviour of $\phi'(r,\lambda)$ is obtained by differentiating (2.1) with respect to r and proceeding as before.) Hence, for large values of $|\lambda|$, the zeros of $\omega(\lambda)$ must be near the zeros of $\sin(b\sqrt{\lambda}-\frac{1}{2}l\pi-\frac{1}{2}\pi)$, which

are, of course, independent of $q(r)$. Further, for large $|\lambda|$, the zeros of $\omega(\lambda)$ are simple. This is best seen by writing

$$\omega'(\lambda) = \frac{1}{2\pi i}\int\limits_C \frac{\omega(\mu)\,d\mu}{(\mu-\lambda)^2},$$

where C is a circle with centre λ, and by using the asymptotic expression (3.1) for $\omega(\lambda)$ to give an asymptotic expression for $\omega'(\lambda)$. It is then clear that values of λ near the zeros of $\sin(b\surd\lambda-\frac{1}{2}l\pi-\frac{1}{2}\pi)$ do not satisfy $\omega'(\lambda) = 0$.

We now construct the function $\Phi(r,\lambda)$, where

$$\Phi(r,\lambda) = \frac{\chi(r,\lambda)}{\omega(\lambda)}\int\limits_0^r \phi(t,\lambda)f(t)\,dt+\frac{\phi(r,\lambda)}{\omega(\lambda)}\int\limits_r^b \chi(t,\lambda)f(t)\,dt,$$

and $f(t)$ is any function which is $L^2(0,b)$. This is a meromorphic function of λ, having poles at the zeros of $\omega(\lambda)$. It will be our object in the next section to show that, if $f(t)$ is such that all the residues of $\Phi(r,\lambda)$ at its poles vanish, then $f(t) = 0$ almost everywhere.

4. If all the residues vanish, $\Phi(r,\lambda)$ becomes an integral function of λ. Let us suppose that we can prove (as we shall do) that we can find a sequence of circles $|\lambda| = R_n$, with $R_n \to \infty$, such that $\Phi(r,\lambda)$ is bounded on the circles, with the bound possibly dependent on r, but independent of n. Then, by Liouville's theorem, $\Phi(r,\lambda)$ is a constant, independent of λ.

Suppose then that $\Phi(r,\lambda) = g(r)$. It follows by differentiation that

$$\frac{d^2g}{dr^2}+\left\{\lambda-q(r)-\frac{l(l+1)}{r^2}\right\}g = f(r),$$

with the result holding at least almost everywhere. By varying λ, we have $g(r) = 0$, and hence $f(r) = 0$ almost everywhere.

It remains to prove the boundedness of $\Phi(r,\lambda)$, with r fixed, but $|\lambda| \to \infty$, on the circles $|\lambda| = R_n$. Since we are concerned only with results 'almost everywhere', we may exclude $r = 0$. The differential equation is thus non-singular in the interval $[r,b]$, and we can appeal to (1) [equation (1.7.8)] to get an asymptotic form of $\chi(r,\lambda)$ for sufficiently large $|\lambda|$. In fact, we have

$$|\chi(r,\lambda)| < Ae^{(b-r)|\tau|}, \tag{4.1}$$

where A denotes various positive constants independent of λ. From § 2 we have, again for fixed r and sufficiently large $|\lambda|$,

$$|\phi(r,\lambda)| < A\omega(r,\lambda). \tag{4.2}$$

Finally, if we choose the sequence $\{R_n\}$ to be such that

$$b\sqrt{R_n}-\tfrac{1}{2}l\pi-\tfrac{1}{2}\pi = (n+\tfrac{1}{2})\pi,$$

we see that
$$\sin(b\sqrt{\lambda}-\tfrac{1}{2}l\pi-\tfrac{1}{2}\pi) > Ae^{b|\tau|}$$

on each of the circles $|\lambda| = R_n$, and so, on those circles, for n sufficiently large, we have from (3.1) that

$$|\omega(\lambda)| > A|\lambda|^{\frac{1}{2}}e^{b|\tau|}. \tag{4.3}$$

If we now substitute (4.1), (4.2), (4.3) in the definition of $\Phi(r,\lambda)$, and use Schwarz's inequality to estimate the integrals, we see readily that, on the circles $|\lambda| = R_n$, $\Phi(r,\lambda)$ is bounded with bound independent of n.

5. From this, we can deduce the completeness of the eigenfunctions and adjoint functions. Before we do this, however, we must examine the nature of these eigenfunctions and adjoint functions. In the real self-adjoint case, it is well known that the zeros of $\omega(\lambda)$ are real and simple, and, if λ_n is such a zero, $\chi(r,\lambda_n)$ is a multiple of $\phi(r,\lambda_n)$, so that we may write $\chi(r,\lambda_n) = k_n\phi(r,\lambda_n)$. Then, near $\lambda = \lambda_n$, the singular part of $\Phi(r,\lambda)$ is

$$\frac{k_n\phi(r,\lambda_n)}{(\lambda-\lambda_n)\omega'(\lambda_n)}\int\limits_0^b \phi(t,\lambda_n)f(t)\,dt.$$

Hence the residue at $\lambda = \lambda_n$ is

$$\frac{k_n\phi(r,\lambda_n)}{\omega'(\lambda_n)}\int\limits_0^b \phi(t,\lambda_n)f(t)\,dt,$$

and this vanishes for all r if and only if the Fourier coefficient of $f(t)$ with respect to the eigenfunction $\phi(t,\lambda_n)$ vanishes.

This argument remains valid even in the non-self-adjoint case provided that λ_n is a simple zero of $\omega(\lambda)$. However, there is no longer any guarantee that the eigenvalues of $\omega(\lambda)$ will be simple, and counter-examples are easily provided.

Suppose now that λ_n is a zero of order p of $\omega(\lambda)$. Then, at $\lambda = \lambda_n$, $\Phi(r,\lambda)$ has a residue of the form

$$\sum_{s=0}^{p-1} A_s(\lambda_n)\left\{\int\limits_0^r \left[\frac{d^s}{d\lambda^s}\{\chi(r,\lambda)\phi(y,\lambda)\}\right]_{\lambda=\lambda_n} f(y)\,dy+\right.$$
$$\left.+\int\limits_r^b \left[\frac{d^s}{d\lambda^s}\{\phi(r,\lambda)\chi(y,\lambda)\}\right]_{\lambda=\lambda_n} f(y)\,dy\right\}, \tag{5.1}$$

where the $A_s(\lambda_n)$ are constants depending on the derivatives of $\omega(\lambda)$ at $\lambda = \lambda_n$ and whose precise value will not concern us.

Now $\omega(\lambda)$ can be written in the form

$$\phi^2(r,\lambda)\frac{d}{dr}\{\chi(r,\lambda)/\phi(r,\lambda)\},$$

and we know that $\omega(\lambda_n) = \omega'(\lambda_n) = 0$. Hence

$$\left\{\frac{d}{d\lambda}\left[\frac{d}{dr}\{\chi(r,\lambda)/\phi(r,\lambda)\}\right]\right\}_{\lambda=\lambda_n} = 0,$$

and interchange of the order of differentiation gives that

$$\left[\frac{d}{d\lambda}\{\chi(r,\lambda)/\phi(r,\lambda)\}\right]_{\lambda=\lambda_n}$$

is independent of r.

If we repeat this process with higher differentiations with respect to λ, we obtain finally that

$$\left[\frac{d^s}{d\lambda^s}\{\chi(r,\lambda)/\phi(r,\lambda)\}\right]_{\lambda=\lambda_n}$$

is independent of r for $s = 0, 1,..., p-1$. This implies that, for these values of s,

$$\begin{aligned}\left[\frac{d^s}{d\lambda^s}\{\chi(r,\lambda)\phi(y,\lambda)\}\right]_{\lambda=\lambda_n} &= \left[\frac{d^s}{d\lambda^s}\{\phi(r,\lambda)[\chi(r,\lambda)/\phi(r,\lambda)]\phi(y,\lambda)\}\right]_{\lambda=\lambda_n} \\ &= \left[\frac{d^s}{d\lambda^s}\{\phi(r,\lambda)[\chi(y,\lambda)/\phi(y,\lambda)]\phi(y,\lambda)\}\right]_{\lambda=\lambda_n} \\ &= \left[\frac{d^s}{d\lambda^s}\{\phi(r,\lambda)\chi(y,\lambda)\}\right]_{\lambda=\lambda_n},\end{aligned}$$

so that (5.1) can be expressed as a linear combination of the p functions $[d^s\{\phi(r,\lambda)\}/d\lambda^s]_{\lambda=\lambda_n}$, the coefficients being homogeneous linear combinations of the p expressions

$$\int\limits_0^b \left[\frac{d^s}{d\lambda^s}\{\chi(y,\lambda)\}\right]_{\lambda=\lambda_n} f(y)\,dy,$$

or, what is the same thing, homogeneous linear combinations of the p expressions

$$\int\limits_0^b \left[\frac{d^s}{d\lambda^s}\{\phi(y,\lambda)\}\right]_{\lambda=\lambda_n} f(y)\,dy.$$

For the residue to vanish it is therefore sufficient that all the Fourier coefficients of $f(t)$ with respect to the p functions $[d^s\{\phi(t,\lambda)\}/d\lambda^s]_{\lambda=\lambda_n}$

should vanish. Hence, if all the Fourier coefficients of $f(t)$ vanish at all zeros of $\omega(\lambda)$, then all the residues of $\Phi(r, \lambda)$ vanish, and so, as already proved, $f(t) = 0$ almost everywhere. This shows, by application of a standard theorem, that the system of eigenfunctions and adjoint functions, where the adjoint functions are

$$\left[\frac{d^s}{d\lambda^s}\{\phi(t, \lambda)\}\right]_{\lambda=\lambda_n} \quad (s = 1,..., p-1),$$

is complete.

The question does arise whether the adjoint functions are indeed necessary for completeness, or whether on the contrary they themselves can be expressed as linear combinations of the eigenfunctions, and so be eliminated from the expansion of an arbitrary function. It is a standard theorem in the theory of orthogonal functions that all the eigenfunctions and adjoint functions are necessary if they form an *orthonormal* set, and we shall prove that they are substantially orthonormal in § 6. What we shall actually prove (and it is clear that this will be sufficient) is

(i) that all the eigenfunctions and adjoint functions associated with an eigenvalue λ_n are orthogonal to all the eigenfunctions and adjoint functions associated with an eigenvalue λ_m, where $n \neq m$;

(ii) that the eigenfunctions and adjoint functions associated with an eigenvalue λ_n of multiplicity p can be expressed by a non-singular transformation as linear combinations of p orthonormal functions.

It should be remarked that the number of multiple eigenvalues is at most finite, and so the number of adjoint functions is at most finite. For we remarked in § 3 that, for $|\lambda|$ sufficiently large, the zeros of $\omega(\lambda)$ are simple.

Finally, it is worth mentioning that, formally, the adjoint functions satisfy differential equations related to the original differential equation (1.4). For, differentiating (1.4) with respect to λ, we have

$$\frac{d^2}{dr^2}\left\{\frac{d\phi(r, \lambda)}{d\lambda}\right\}+\left\{\lambda-q(r)-\frac{l(l+1)}{r^2}\right\}\frac{d\phi(r, \lambda)}{d\lambda} = -\phi(r, \lambda),$$

so that $d\phi(r, \lambda)/d\lambda$ satisfies the equation

$$\left\{\frac{d^2}{dr^2}+\lambda-q(r)-\frac{l(l+1)}{r^2}\right\}^2 y = 0.$$

Similarly $d^s\phi(r, \lambda)/d\lambda^s$ satisfies the equation

$$\left\{\frac{d^2}{dr^2}+\lambda-q(r)-\frac{l(l+1)}{r^2}\right\}^{s+1} y = 0.$$

This ties up with the way in which the adjoint functions appear in the general operator approach (9).

6. The required orthonormality results are obtained in four stages.

(I) We discuss the boundary conditions on the adjoint functions at $r = 0, b$.

(II) We prove the result (i) of § 5.

(III) We prove that, if λ_n is an eigenvalue of multiplicity p and

$$\phi_i = \left[\frac{d^i\phi(r,\lambda)}{d\lambda^i}\right]_{\lambda=\lambda_n} \quad (i = 0, 1, 2, \ldots),$$

then ϕ_i is orthogonal to ϕ_j if $i+j \leqslant p-2$, but ϕ_i is not orthogonal to ϕ_j if $i+j = p-1$.

(IV) We prove finally that there is a non-singular transformation from the eigenfunctions and adjoint functions ϕ_i to an orthonormal set.

(I) If $\omega(\lambda)$ has a zero of order p at $\lambda = \lambda_n$, then

$$\omega(\lambda) = -\cos\beta\phi(b,\lambda) - \sin\beta\phi'(b,\lambda) = 0 \quad \text{at } \lambda = \lambda_n,$$

$$\omega'(\lambda) = -\cos\beta\frac{d\phi(b,\lambda)}{d\lambda} - \sin\beta\left[\frac{d}{dr}\frac{d\phi(r,\lambda)}{d\lambda}\right]_{r=b} = 0 \quad \text{at } \lambda = \lambda_n,$$

$$\cdots\cdots\cdots\cdots$$

$$\omega^{(p-1)}(\lambda) = -\cos\beta\frac{d^{p-1}\phi(b,\lambda)}{d\lambda^{p-1}} - \sin\beta\left[\frac{d}{dr}\frac{d^{p-1}\phi(r,\lambda)}{d\lambda^{p-1}}\right]_{r=b} = 0 \quad \text{at } \lambda = \lambda_n.$$

Consequently, $\phi_0, \phi_1, \ldots, \phi_{p-1}$ all satisfy the same homogeneous boundary condition at $r = b$, while at $r = 0$, since the boundary conditions on $\phi(r,\lambda)$ are independent of λ,

$$\phi_1 = 0 = \frac{d\phi_1}{dr}, \quad \ldots, \quad \phi_{p-1} = 0 = \frac{d\phi_{p-1}}{dr}.$$

(II) Suppose that $\omega(\lambda)$ has a zero of order p at $\lambda = \lambda_n$ and of order s at $\lambda = \lambda_m$. Let us put

$$\psi_k = \left[\frac{d^k\phi(r,\lambda)}{d\lambda^k}\right]_{\lambda=\lambda_m} \quad (k = 0, 1, 2, \ldots).$$

We shall prove ϕ_i, ψ_j orthogonal by induction, where $i = 0, 1, \ldots, p-1$; $j = 0, 1, \ldots, s-1$.

We shall assume the orthogonality when $i+j = q$, and prove it when $i+j = q+1$. (The proof of the result when $q = 0$, which is necessary to complete the induction proof, goes through as in the real case.)

Differentiating (1.4) i times with respect to λ and putting $\lambda = \lambda_n$, we have

$$\frac{d^2\phi_i}{dr^2} + \left\{\lambda_n - q(r) - \frac{l(l+1)}{r^2}\right\}\phi_i + i\phi_{i-1} = 0;$$

and similarly

$$\frac{d^2\psi_j}{dr^2}+\left\{\lambda_m-q(r)-\frac{l(l+1)}{r^2}\right\}\psi_j+j\psi_{j-1}=0.$$

We multiply the first equation by ψ_j, the second by ϕ_i, subtract and integrate. Using the boundary conditions obtained in (I), we have

$$(\lambda_n-\lambda_m)\int\limits_0^b \phi_i\psi_j\,dr+i\int\limits_0^b \phi_{i-1}\psi_j\,dr-j\int\limits_0^b \phi_i\psi_{j-1}\,dr=0,$$

so that, by the induction hypothesis,

$$\int\limits_0^b \phi_i\psi_j\,dr=0.$$

(III) An induction proof also gives the orthogonality of ϕ_i to ϕ_j $(i+j\leqslant p-2)$. We assume for $i\leqslant q$ that ϕ_i is orthogonal to all ϕ_j such that $i+j\leqslant p-2$, and then prove this true for $i=q+1$. Finally, we prove it for $i=0$. For writing down the differential equations for ϕ_{j+1}, ϕ_{q+1}, we obtain by the usual process that

$$(j+1)\int\limits_0^b \phi_j\phi_{q+1}\,dr-(q+1)\int\limits_0^b \phi_{j+1}\phi_q\,dr=0,$$

whence, by the induction hypothesis,

$$\int\limits_0^b \phi_j\phi_{q+1}\,dr=0.$$

The result is proved for $i=0$ by carrying through the above argument with $q+1=0$.

We can further prove that

$$\int\limits_0^b \phi_i\phi_j\,dr\neq 0$$

if $i+j=p-1$. For $\omega^{(p)}(\lambda_n)\neq 0$, and so

$$\left[\phi_0\frac{d\phi_p}{dr}-\phi_p\frac{d\phi_0}{dr}\right]_{r=b}\neq 0.$$

Hence writing down the equations for ϕ_0, ϕ_p, we obtain by the usual process

$$\int\limits_0^b \phi_0\phi_{p-1}\,dr\neq 0.$$

Then, with ϕ_1, ϕ_{p-1}, we have

$$(p-1)\int_0^b \phi_1\phi_{p-2}\,dr-\int_0^b \phi_0\phi_{p-1}\,dr = 0,$$

so that

$$\int_0^b \phi_1\phi_{p-2}\,dr \neq 0.$$

This can be continued to give the required result.

(IV) Let us make the transformation

$$\boldsymbol{\theta} = C\boldsymbol{\varphi},$$

where $\boldsymbol{\varphi}$ is the column vector $[\phi_i]$ and C is a $p\times p$ matrix whose elements may be complex but are independent of r and λ. Then, if $\boldsymbol{\theta} = [\theta_i]$, $[\theta_i\theta_j] = \boldsymbol{\theta\theta}' = C\boldsymbol{\varphi\varphi}'C' = C[\phi_i\phi_j]C'$, and so

$$\left[\int_0^b \theta_i\theta_j\,dr\right] = C\left[\int_0^b \phi_i\phi_j\,dr\right]C'.$$

Now the matrix $\left[\int_0^b \phi_i\phi_j\,dr\right]$ is non-singular, for from (III) each element above the *top-right* to *bottom-left* diagonal is zero, and each element on the diagonal is non-zero. Hence it is possible to find a matrix C such that

$$C\left[\int_0^b \phi_i\phi_j\,dr\right]C' = I,$$

where I is the unit matrix, and this gives the required transformation to an orthonormal set.

REFERENCES

1. E. C. Titchmarsh, *Eigenfunction expansions associated with second-order differential equations* (Oxford, 1946).
2. G. D. Birkhoff, 'Boundary value and expansion problems of ordinary linear differential equations', *Trans. American Math. Soc.* 9 (1908) 373–95.
3. J. Tamarkin, 'Sur quelques points de la théorie des équations différentielles', *Rend. Circ. mat. Palermo* 34 (1912) 345–82.
4. G. D. Birkhoff, 'Note on the expansion problem of ordinary linear differential equations', ibid. 36 (1913) 115–26.
5. J. Tamarkin, 'Sur un problème de la théorie des équations différentielles linéaires ordinaires', ibid. 37 (1914) 376–8.
6. —— 'Some general problems of the theory of ordinary linear differential equations and expansion of an arbitrary function in series of fundamental functions', *Math. Zeit.* 27 (1927) 1–54.

7. R. Peierls, 'Expansions in terms of sets of functions with complex eigenvalues', *Proc. Cambridge Phil. Soc.* **44** (1948) 242–50.
8. C. L. Dolph, 'Recent developments in some non-self-adjoint problems of mathematical physics', *Bull. American Math. Soc.* **67** (1961) 1–69.
9. J. Schwartz, 'Perturbations of spectral operators, and applications', *Pacific J. of Math.* 4 (1954) 415–58.
10. M. V. Keldysh, 'On the characteristic values and characteristic functions of certain classes of non-self-adjoint equations', *Dokl. Akad. Nauk SSSR* (n.s.) 77 (1951) 11–14.

EQUATIONS OF THE FORM $f(x) = g(y)$

By H. DAVENPORT, D. J. LEWIS, *and* A. SCHINZEL (*Cambridge*)

[Received 6 August 1961]

1. In this note we obtain some general conditions which will ensure that an equation of the form $f(x) = g(y)$, where f, g are given polynomials with integral coefficients, has at most a finite number of integral solutions. We apply the result to the particular case

$$f(x) = x^n+x^{n-1}+...+x, \qquad g(y) = y^m+y^{m-1}+...+y, \tag{1}$$

which was the subject of a recent note by Mąkowski and Schinzel.†

A theoretically complete answer to the question whether a Diophantine equation $f(x,y) = 0$ has infinitely many integral solutions or not was given by the work of Siegel.‡ Provided that $f(x,y)$ is irreducible over the complex number field, there are infinitely many solutions if and only if there is a rational parametric solution of a special form; and in particular there are at most finitely many solutions if the genus is positive. The requirement of irreducibility results from the fact that the notion of genus, which plays an important part in Siegel's work, applies only to irreducible equations.§

Thus the first question to decide is whether $f(x)-g(y)$ is irreducible over the complex field, and it seems to us that this question is of interest in itself. There is one obvious case of reducibility, namely when

$$f(x) = F(f_1(x)), \qquad g(y) = F(g_1(y)), \tag{2}$$

where F, f_1, g_1 are arbitrary polynomials, subject to $\deg F > 1$. We have found a further case, namely

$$f(x) = cF_k(f_1(x)), \qquad g(y) = -cF_k(g_1(y)), \tag{3}$$

where c is a constant, f_1, g_1 are arbitrary polynomials, and F_k is the special polynomial defined by

$$2F_k(z) = \{z+\sqrt{(z^2-1)}\}^k+\{z-\sqrt{(z^2-1)}\}^k, \tag{4}$$

k being an even positive integer. [If k were odd, we should have

† 'Sur l'équation indéterminée de M. Goormaghtigh', *Mathesis* 68 (1959) 128–42.

‡ For a brief account and references, see Skolem, *Diophantische Gleichungen* (Ergebnisse der Math. v. 4) 102–4.

§ Irreducibility is not needed if one can prove directly that an equation does not admit a rational parametric solution of the special form, but this does not seem to be easy for the equations under consideration.

Quart. J. Math. Oxford (2), 12 (1961), 304-12.

$-F_k(g_1(y)) = F_k(-g_1(y))$, and (3) would be an instance of (2).] To obtain the factorization, we note that

$$F_k(\cosh\theta) = \cosh k\theta.$$

Hence, if $u = \cosh\alpha$ and $v = \cosh\beta$, we have

$$F_k(u)+F_k(v) = \cosh k\alpha+\cosh k\beta.$$

This vanishes if $\alpha = \beta+r\pi i/k$, where r is an odd integer, and therefore

$$\cosh\alpha-\cosh\beta\cos\frac{r\pi}{k}-i\sinh\beta\sin\frac{r\pi}{k}$$

is a factor. Taking conjugate complex factors together, we find that

$$F_k(u)+F_k(v) = C\prod_{\substack{r=1\\ r\,\mathrm{odd}}}^{k-1}\left(u^2-2uv\cos\frac{r\pi}{k}+v^2-\sin^2\frac{r\pi}{k}\right), \tag{5}$$

where C is a constant. So $F_k(f_1(x))+F_k(g_1(y))$ factorizes if $k > 2$.

We can suppose not only that k is even but that k is a power of 2; for, if $k = 2^s t$, where $t > 1$ is odd, we have $F_k(u) = F_t(U)$, where $U = F_{2^s}(u)$. Since $-F_t(U) = F_t(-U)$, we have again an instance of (2).

When k is a power of 2, the polynomials $f(x)$, $g(y)$ in (3) cannot in general be expressed in the form (2). For suppose that

$$F_k(f_1(x)) = F(f_2(x)), \qquad F_k(g_1(y)) = -F(g_2(y)).$$

If this holds identically in $f_1(x)$ it will hold for $f_1(x) = x$, whence

$$F_{2^\kappa}(x) = F(f_3(x)), \qquad F_{2^\kappa}(x) = -F(g_3(x)).$$

Hence $f_3(u)-g_3(v)$ divides (5), and therefore the highest coefficients in $f_3(u)$, $g_3(v)$ must have a real ratio. As the degree of F is a power of 2 (greater than 1), the two identities last stated contradict one another on making $x \to \infty$.

It may be observed that it suffices to take $k = 4$ in (3), since $F_{2^{s+4}}(u) = F_4(U)$ in the notation used above. It follows that (3) can be written as

$$f(x) = 8X^4-8X^2+1, \qquad g(y) = -8Y^4+8Y^2-1,$$

where X, Y are polynomials in x, y respectively, and the factorization is

$$f(x)-g(y) = 2\{2X^2+2\sqrt{2}XY+2Y^2-1\}\{2X^2-2\sqrt{2}XY+2Y^2-1\}.$$

We have not been able to settle the question whether $f(x)-g(y)$ can be reducible in any cases other than those in (2) and (3). We give a simple condition which will ensure irreducibility, and it so happens that this condition also ensures positive genus, except in two particular cases.

Let $f(x)$ be of degree $n > 1$ and $g(y)$ of degree $m > 1$. Let

$$D(\lambda) = \operatorname{disc}(f(x)+\lambda), \qquad E(\lambda) = \operatorname{disc}(g(y)+\lambda); \tag{6}$$

these are polynomials in λ of degrees $n-1$, $m-1$ respectively.

THEOREM 1. *Suppose there are at least $[\frac{1}{2}n]$ distinct roots of $D(\lambda) = 0$ for which $E(\lambda) \neq 0$. Then $f(x)-g(y)$ is irreducible over the complex field. Further, the genus of the equation $f(x)-g(y) = 0$ is strictly positive except possibly when $m = 2$ or $m = n = 3$. Apart from these possible exceptions, the equation has at most a finite number of integral solutions.*

In § 4 we apply this to the special equation with f, g as in (1) and prove (Theorem 2) that there are at most finitely many solutions if $m > 1$, $n > 1$, $m \neq n$.

In § 5 we discuss two cases in which Runge's method is applicable to this special equation, namely when $(m, n) > 1$ or when $n = m+1$. This method allows one, in any particular case, to calculate an upper bound for any possible solution.

2. LEMMA 1. *Suppose that*

$$f(x)-g(y) = \phi(x,y)\psi(x,y) \tag{7}$$

identically, where ϕ, ψ are non-constant polynomials with real or complex coefficients. Suppose $D(\lambda_1) = 0$, $E(\lambda_1) \neq 0$, and let x_1 satisfy

$$f(x_1)+\lambda_1 = 0, \qquad f'(x_1) = 0. \tag{8}$$

Then $x-x_1$ divides identically both $\phi'(x,y)$ and $\psi'(x,y)$, where

$$\phi' = \partial\phi/\partial x, \qquad \psi' = \partial\psi/\partial x.$$

Proof. The existence of x_1 satisfying (8) follows from $D(\lambda_1) = 0$. Since $E(\lambda_1) \neq 0$, there are m distinct values of y satisfying $g(y)+\lambda_1 = 0$. Since

$$-g(y)-\lambda_1 = f(x_1)-g(y) = \phi(x_1,y)\psi(x_1,y),$$

each of these m values of y must satisfy either

$$\phi(x_1,y) = 0 \quad \text{or} \quad \psi(x_1,y) = 0.$$

By the identity (7), the highest terms in y in $\phi(x,y)$ and $\psi(x,y)$ do not involve x. Let

$$\deg_y \phi(x,y) = s, \qquad \deg_y \psi(x,y) = m-s.$$

Then s of the m values of y mentioned above constitute the roots of $\phi(x_1,y) = 0$ and the remaining $m-s$ constitute the roots of $\psi(x_1,y) = 0$.

Differentiating (7) partially with respect to x and putting $x = x_1$, we obtain

$$\phi(x_1,y)\psi'(x_1,y)+\psi(x_1,y)\phi'(x_1,y) = 0$$

identically in y. It follows that the s distinct values of y mentioned above satisfy

$$\phi(x_1, y) = 0, \qquad \phi'(x_1, y) = 0,$$

and the remaining $m-s$ satisfy the corresponding equations with ψ for ϕ.

Since the highest term in y in $\phi(x,y)$ does not involve x, we have

$$\deg_y \phi'(x,y) \leqslant s-1.$$

Since $\phi'(x_1, y)$ vanishes for s distinct values of y, it must vanish identically. Hence $x-x_1$ divides $\phi'(x,y)$ identically, and similarly divides $\psi'(x,y)$ identically.

Lemma 2. *Under the hypotheses of Theorem* 1, *$f(x)-g(y)$ is irreducible over the complex field.*

Proof. Let $\lambda_1,..., \lambda_h$, where $h \geqslant [\frac{1}{2}n]$, be distinct roots of $D(\lambda) = 0$ for which $E(\lambda) \neq 0$, and let $x_1,..., x_h$ be determined as in (8). Then $x_1,..., x_h$ are distinct since $f(x_j) = -\lambda_j$. By Lemma 1, if (7) holds, then $\phi'(x,y)$ is divisible identically by $(x-x_1)...(x-x_h)$.

Let $r = \deg_x \phi(x,y)$; actually $r/s = n/m$, but we do not need this fact. By interchanging ϕ, ψ if necessary we can suppose that $r \leqslant [\frac{1}{2}n]$. We now have a contradiction, since

$$\deg_x \phi'(x,y) \leqslant r-1,$$

so that $\phi'(x,y)$ cannot have a factor of degree $h \geqslant [\frac{1}{2}n] \geqslant r$ in x. This proves the result.

It may be remarked that, if the hypothesis of Theorem 1 were relaxed, the irreducibility of $f(x)-g(y)$ might fail. For, if

$$f(x) = x^2(x^k-1)^2, \qquad g(y) = y^2,$$

where k is odd, it is easily seen that $D(\lambda) = 0$ has k roots

$$\lambda = -k^2(k+1)^{-2-2/k}$$

other than $\lambda = 0$; thus there are $[\frac{1}{2}n]-1$ distinct roots of $D(\lambda) = 0$ which are not roots of $E(\lambda) = 0$, but $f(x)-g(y)$ is reducible.

Lemma 3. *Under the hypotheses of Theorem* 1, *the genus of the equation $f(x)-g(y) = 0$ is positive, except possibly if $m = 2$ or $n = m = 3$.*

Proof. We consider the equation as defining x as an algebraic function of y. In the neighbourhood of each point y in the complex plane for which there is a multiple solution in x (i.e. each branch point), the values of x fall into cycles, those in each cycle undergoing a cyclic permutation as y goes round the point. There may also be cycles in

the neighbourhood of $y = \infty$, found by putting $y = 1/y'$. The genus g is given by†

$$g = \tfrac{1}{2}\sum (r-1) - n + 1, \tag{9}$$

where the summation is over all cycles, and r is the number of values of x in a particular cycle. If we omit any cycles from the sum, we obtain a lower bound for g.

The finite branch points are given by $g(y) = -\lambda$, where $D(\lambda) = 0$. We consider only those roots $\lambda_1,\ldots, \lambda_h$ ($h \geqslant [\frac{1}{2}n]$) for which $E(\lambda_j) \neq 0$. To each λ_j there correspond m distinct branch points $y_1^{(j)},\ldots, y_m^{(j)}$, and all these mh branch points are distinct. At each of these there is at least one cycle with $r \geqslant 2$. The contribution of these branch points to $\frac{1}{2}\sum (r-1)$ is at least $\frac{1}{2}m[\frac{1}{2}n]$.

To investigate the point at infinity, we put $y = 1/y'$ and $x = 1/x'$. The equation becomes

$$y'^m f_0(x') - x'^n g_0(y') = 0,$$

where f_0, g_0 are polynomials with $f_0(0) \neq 0$, $g_0(0) \neq 0$. The values of x' near $y' = 0$ are given by n power series in y'^{1/n_1}, where $n_1 = n/(m,n)$. These values fall into (m,n) cycles, each containing $n/(m,n)$ values. Thus the contribution to $\frac{1}{2}\sum (r-1)$ is $\frac{1}{2}(n-(m,n))$.

Finally,

$$g \geqslant \tfrac{1}{2}m[\tfrac{1}{2}n] + \tfrac{1}{2}(n-(n,m)) - n + 1 = \tfrac{1}{2}m[\tfrac{1}{2}n] - \tfrac{1}{2}n - \tfrac{1}{2}(n,m) + 1.$$

Plainly $g > 0$ if $m \geqslant 6$ and $n > 1$, and a simple enumeration of cases shows that $g > 0$ when $m > 2$ and $n > 1$, except when $m = n = 3$.

We add the remark that the two exceptional cases $m = 2$ and $m = n = 3$ are genuine. First, if

$$f(x) = x(x^k-1)^2, \qquad g(y) = y^2, \tag{10}$$

we find that the equation $D(\lambda) = 0$ has $k = [\frac{1}{2}n]$ roots other than 0, namely those given by

$$(-\lambda)^k = (2k)^{2k}(2k+1)^{-2k-1}.$$

Thus the hypothesis is satisfied, but the genus is 0. Secondly, if

$$f(x) = x^2(x+1), \qquad g(y) = y^2, \tag{11}$$

then $D(\lambda) = 0$ has one root $\lambda = -4/27$ other than 0, and again the hypothesis is satisfied but $g = 0$.

Theorem 1 follows from Lemmas 2, 3 and the work of Siegel.

3. We have supposed in Theorem 1, for simplicity, that there are at least $[\frac{1}{2}n]$ *distinct* roots of $D(\lambda) = 0$ for which $E(\lambda) \neq 0$. But the

† G. A. Bliss, *Algebraic functions* (A.M.S. Colloquium Publications 16, 1933) Theorem 31.6.

conclusions still hold if the number of roots of $D(\lambda) = 0$ which are not roots of $E(\lambda) = 0$, *counted with their multiplicities*, is at least $[\frac{1}{2}n]$. We indicate briefly the necessary modifications to the proof.

Suppose λ_1 a root of $D(\lambda) = 0$ of multiplicity e_1. Let $x_1^{(1)}, x_1^{(2)},\ldots$ be the solutions of (8), and let their multiplicities as solutions of $f'(x) = 0$ be $e_1^{(1)}, e_1^{(2)},\ldots$, so that their multiplicities as solutions of $f(x)+\lambda_1 = 0$ are $e_1^{(1)}+1$, $e_1^{(2)}+1,\ldots$. Since

$$D(\lambda) = C \prod_{f'(x)=0} (f(x)+\lambda),$$

where C is a numerical constant, we have $e_1 = e_1^{(1)}+e_1^{(2)}+\ldots$. By a slight modification of the proof of Lemma 1, using higher partial derivatives, we find that $\phi'(x,y)$ is divisible identically by

$$(x-x_1^{(1)})^{e_1^{(1)}}(x-x_1^{(2)})^{e_1^{(2)}}\ldots,$$

i.e. by a polynomial in x of degree e_1. Combining the results for the various values of λ we get a contradiction as in the proof of Lemma 2.

In the proof of Lemma 3, we find that in the present more general situation, at each of the m branch points y_j corresponding to λ_1, there are cycles of $e_1^{(1)}+1$ values, of $e_1^{(2)}+1$ values, and so on. Thus the contribution of the finite branch points to the formula for g is still $\geqslant \frac{1}{2}m[\frac{1}{2}n]$.

4. We now apply Theorem 1 to the particular case (1), and suppose that $n > m > 1$.

LEMMA 4. *The discriminant $D_n(\lambda)$ of $f(x)+\lambda$, where $f(x)$ is defined in* (1), *is given (apart from a numerical factor) by*

$$D_n(\lambda) = \frac{(\lambda-1)^{n+1}+c_n\lambda^n}{(\lambda+n)^2}, \qquad c_n = \frac{(n+1)^{n+1}}{n^n}. \tag{12}$$

The equation $D_n(\lambda) = 0$ has distinct roots.

Proof. Suppose that the equations $f(x)+\lambda = 0$, $f'(x) = 0$ have a common solution. Then

$$x^{n+1}+(\lambda-1)x-\lambda = 0, \qquad (n+1)x^n+\lambda-1 = 0. \tag{13}$$

These imply
$$n(\lambda-1)x-(n+1)\lambda = 0, \tag{14}$$

and on substituting in $(13)_2$ we get

$$(\lambda-1)^{n+1}+c_n\lambda^n = 0. \tag{15}$$

Now $(\lambda+n)^2$ is a factor of the left-hand side, but for $\lambda = -n$ the equations do not have a common solution since (14) gives $x = 1$ and

$$f'(1) = n+(n-1)+\ldots+1 \neq 0.$$

Hence, if $D_n(\lambda)$ is defined by (12), the condition $D_n(\lambda) = 0$ is necessary

for a common solution. It is also sufficient, the common solution being given by (14).

The roots of $D_n(\lambda) = 0$ are distinct since (15) and its derivative imply $\lambda = -n$, and $\lambda = -n$ does not satisfy $D_n(\lambda) = 0$. Since $D_n(\lambda)$ is of the correct degree $n-1$ and has no repeated factor, it is the discriminant of $f(x)+\lambda$, except for a numerical factor.

THEOREM 2. *If $n > m > 1$ and $f(x)$, $g(y)$ are defined by* (1), *the equation $f(x) = g(y)$ has at most a finite number of integral solutions.*

Proof. We prove that the hypothesis of Theorem 1 is satisfied. A common root of $D_n(\lambda) = 0$ and $D_m(\lambda) = 0$ satisfies

$$(\lambda-1)^{n+1}+c_n\lambda^n = 0, \qquad (\lambda-1)^{m+1}+c_m\lambda^m = 0,$$

whence
$$c_m(\lambda-1)^{n-m}-c_n\lambda^{n-m} = 0.$$

Hence the number of common roots is at most $n-m$. It is also at most $m-1$, since this is the degree of $D_m(\lambda)$, and so it is

$$\leqslant \min(n-m, m-1) \leqslant n-1-[\tfrac{1}{2}n].$$

Thus the number of roots of $D_n(\lambda)$ for which $D_m(\lambda) \neq 0$ is at least $[\tfrac{1}{2}n]$.

The desired conclusion now follows from Theorem 1, except possibly when $m = 2$. The equation is now

$$y^2+y = x^n+x^{n-1}+...+x,$$

i.e.
$$(y+\tfrac{1}{2})^2 = x^n+x^{n-1}+...+x+\tfrac{1}{4},$$

where $n > 2$. This has genus g if $n = 2g+1$ or $2g+2$, provided that the right-hand side has no repeated factor. This is so if $D_n(\tfrac{1}{4}) \neq 0$: that is, if

$$(-\tfrac{3}{4})^{n+1}n^n+(\tfrac{1}{4})^n(n+1)^{n+1} \neq 0,$$

which is obviously true for $n > 2$.

5. Runge's method, when applicable, enables one in principle to give an upper bound for the size of any solution of $f(x,y) = 0$, and so to enumerate the solutions in any numerical case.† Assuming $f(x,y)$ irreducible over the rational field, the method is applicable except when

(i) the highest terms in x and y occur separately as ax^n, by^m;

(ii) each branch of the algebraic function y of x defined by $f = 0$ tends to infinity with x and is of order $x^{n/m}$;

(iii) every term $x^\alpha y^\beta$ in f has $\alpha m+\beta n \leqslant mn$, and the sum of the terms with $\alpha m+\beta n = mn$ is expressible as

$$C(x^{n_1}-\theta_1 y^{m_1})...(x^{n_1}-\theta_k y^{m_1}), \tag{16}$$

where $k = n/n_1 = m/m_1$, and $\prod (X-\theta_j Y)$ is a power of a rationally

† For an account and references, see Skolem, loc. cit. 89–91.

irreducible polynomial (possibly the first power), and at least one θ_j is real.

The method is not applicable to the equation

$$x^n+x^{n-1}+...+x = y^m+y^{m-1}+...+y \quad (n > m > 1) \tag{17}$$

as it stands, unless $(n, m) = d > 1$. In the latter case, (16) becomes

$$\prod_{\zeta} (x^{n_1}-\zeta y^{m_1}),$$

where ζ runs through the dth roots of unity and $n_1 = n/d$, $m_1 = m/d$. The polynomial $\prod (X-\zeta Y)$ is not a power of a rationally irreducible polynomial since it has the factor $X-Y$, and so the method is applicable.

If $n = m+1$, we can make a preliminary transformation, after which Runge's method is applicable in a slightly modified form. We can write the equation as

$$x^{m+1} = (y-x)\{1+(x+y)+...+(x^{m-1}+...+y^{m-1})\}. \tag{18}$$

Since the two factors on the right are relatively prime, we have

$$y-x = \epsilon u^{m+1}, \qquad 1+(x+y)+...+(x^{m-1}+...+y^{m-1}) = \epsilon v^{m+1},$$

where $\epsilon = \pm 1$, and $x = uv$, whence

$$1+u(2v+\epsilon u^m)+...+u^{m-1}\{(v+\epsilon u^m)^{m-1}+...+v^{m-1}\}-\epsilon v^{m+1} = 0. \tag{19}$$

The highest term in v is $-\epsilon v^{m+1}$ and in u is $\epsilon^{m-1}u^{(m+1)(m-1)}$, and the sum of the terms considered in (iii) above is simply

$$\epsilon^{m-1}u^{(m+1)(m-1)}-\epsilon v^{m+1}.$$

We do not know that the left-hand side of (19), say $Q(u, v)$, is rationally irreducible, so we put

$$Q(u,v) = Q_1(u,v)... Q_l(u,v),$$

where the polynomials on the right are rationally irreducible. The product of the highest isobaric parts of the Q_j, giving u the weight 1 and v the weight $m-1$, is

$$R_1(u,v)... R_l(u,v) = \epsilon^{m-1}u^{(m+1)(m-1)}-\epsilon v^{m+1}$$
$$= \epsilon^{m-1} \prod (u^{m-1}-\zeta v),$$

where ζ runs through the $(m+1)$th roots of ϵ^m.

Those equations $Q_j(u,v) = 0$ for which $R_j(u,v)$ contains only factors with complex ζ are amenable to Runge's method. If R_j is divisible by one or both of $u^{m-1}\pm v$, then R_j is rationally reducible unless

$$R_j = u^{m-1}\pm v.$$

In this case Q_j is linear in v, say

$$Q_j(u,v) = P(u)-v,$$

where $P(u)$ is a polynomial in u. But then (18) has the parametric solution

$$x = uP(u), \qquad y = uP(u)+\epsilon u^{m+1},$$

contrary to the fact that (18) has been proved to have positive genus.

Hence Runge's method is applicable to each of the equations

$$Q_j(u, v) = 0.$$

INTERSECTION THEOREMS FOR SYSTEMS OF FINITE SETS

By P. ERDŐS, CHAO KO (*Szechuan*), *and* R. RADO (*Reading*)

[Received 13 August 1961]

1. Introduction

E. SPERNER (1) has proved that *every system of subsets a_ν of a set of finite cardinal m, such that $a_\mu \not\subset a_\nu$ for $\mu \neq \nu$ contains at most $\binom{m}{p}$ elements, where $p = [\frac{1}{2}m]$*. This note concerns analogues of this result. We shall impose an upper limitation on the cardinals of the a_ν and a lower limitation on the cardinals of the intersection of any two sets a_ν, and we shall deduce upper estimates, in many cases best-possible, for the number of elements of such a system of sets a_ν.

2. Notation

The letters a, b, c, d, x, y, z denote finite sets of non-negative integers, all other lower-case letters denote non-negative integers. If $k \leqslant l$, then $[k, l)$ denotes the set

$$\{k, k+1, k+2, ..., l-1\} = \{t\colon k \leqslant t < l\}.$$

The *obliteration operator* $\hat{}$ serves to remove from any system of elements the element above which it is placed. Thus $[k, l) = \{k, k+1, ..., \hat{l}\}$. The *cardinal* of a is $|a|$; *inclusion* (in the wide sense), *union*, *difference*, and *intersection* of sets are denoted by $a \subset b$, $a+b$, $a-b$, ab respectively, and $a-b = a-ab$ for all a, b.

By $S(k, l, m)$ we denote the set of all systems $(a_0, a_1, ..., \hat{a}_n)$ such that

$$a_\nu \subset [0, m);\ |a_\nu| \leqslant l \quad (\nu < n),$$

$$a_\mu \not\subset a_\nu \not\subset a_\mu;\ |a_\mu a_\nu| \geqslant k \quad (\mu < \nu < n).$$

3. Results

THEOREM 1. *If $1 \leqslant l \leqslant \frac{1}{2}m$; $(a_0, ..., \hat{a}_n) \in S(1, l, m)$, then $n \leqslant \binom{m-1}{l-1}$. If, in addition, $|a_\nu| < l$ for some ν, then $n < \binom{m-1}{l-1}$.*

THEOREM 2. *Let $k \leqslant l \leqslant m$, $n \geqslant 2$, $(a_0, ..., \hat{a}_n) \in S(k, l, m)$. Suppose that either*

$$2l \leqslant k+m, \qquad |a_\nu| = l \quad (\nu < n) \tag{1}$$

or†

$$2l \leqslant 1+m, \qquad |a_\nu| \leqslant l \quad (\nu < n). \tag{2}$$

† The condition $|a_\nu| \leqslant l$ is in fact implied by $(a_0, ..., \hat{a}_n) \in S(k, l, m)$.

Quart. J. Math. Oxford (2), 12 (1961), 313–20.

Then (*a*) *either* (i)

$$|a_0 \dots \hat{a}_n| \geqslant k, \qquad n \leqslant \binom{m-k}{l-k},$$

or (ii) $$|a_0 \dots \hat{a}_n| < k < l < m, \qquad n \leqslant \binom{m-k-1}{l-k-1}\binom{l}{k}^3;$$

(*b*) *if* $$m \geqslant k+(l-k)\binom{l}{k}^3, \quad \textit{then} \quad n \leqslant \binom{m-k}{l-k}.$$

Remark. Obviously, if $|a_\nu| = l$ for $\nu < n$, then the upper estimates for n in Theorem 1 and in Theorem 2 (*a*) (i) and (*b*) are best-possible. For, if $k \leqslant l \leqslant m$ and if $a_0, \dots, \hat{a}_n$ are the distinct sets a such that

$$[0,k) \subset a \subset [0,m), \qquad |a| = l,$$

then $$(a_0, \dots, \hat{a}_n) \in S(k,l,m), \qquad n = \binom{m-k}{l-k}.$$

4. The following lemma is due to Sperner (**1**). We give the proof since it is extremely short.

LEMMA. *If*

$$n_0 \geqslant 1, \quad a_\nu \subset [0,m), \quad |a_\nu| = l_0 \quad (\nu < n_0),$$

then there are at least $n_0(m-l_0)(l_0+1)^{-1}$ *sets* b *such that, for some* ν,

$$\nu < n_0, \quad a_\nu \subset b \subset [0,m), \quad |b| = l_0+1. \tag{3}$$

Proof. Let n_1 be the number of sets b defined above. Then, by counting in two different ways the number of pairs (ν, b) satisfying (3), we obtain $n_0(m-l_0) \leqslant n_1(l_0+1)$, which proves the lemma.

5. Proof of Theorem 1

Case 1. Let $|a_\nu| = l\ (\nu < n)$. We have $m \geqslant 2$. If $m = 2$, then $l = 1$; $n = 1 \leqslant \binom{m-1}{l-1}$. Now let $m \geqslant 3$ and use induction over m. Choose, for fixed l, m, n, the a_ν in such a way that the hypothesis holds and, in addition, the number

$$f(a_0, \dots, \hat{a}_n) = s_0 + \dots + \hat{s}_n$$

is minimal, where s_ν is the sum of the elements of a_ν. Put $A = \{a_\nu : \nu < n\}$. If $2l = m$, then $[0,m) - a_\nu \notin A$ and, hence

$$n \leqslant \frac{1}{2}\binom{m}{l} = \binom{m-1}{l-1}.$$

Now let $2l < m$.

Case 1 *a*. Suppose that whenever
$$m-1 \in a \in A, \qquad \lambda \in [0,m)-a,$$
then
$$a-\{m-1\}+\{\lambda\} \in A.$$
We may assume that, for some $n_0 < n$,
$$m-1 \in a_\nu \quad (\nu < n_0), \qquad m-1 \notin a_\nu \quad (n_0 \leqslant \nu < n).$$
Put
$$b_\nu = a_\nu-\{m-1\} \quad (\nu < n_0).$$
Let $\mu < \nu < n_0$. Then
$$|a_\mu+a_\nu| < 2l < m,$$
and there is
$$\lambda \in [0,m)-a_\mu-a_\nu.$$
Then $\quad b_\mu+\{\lambda\} \in A, \qquad b_\mu b_\nu = (b_\mu+\{\lambda\})b_\nu = (b_\mu+\{\lambda\})a_\nu \neq \emptyset$

and therefore
$$l-1 \geqslant 1, \qquad (b_0,\ldots,\hat{b}_{n_0}) \in S(1,l-1,m-1).$$
Since $2(l-1) < m-2 < m-1$ we obtain, by the induction hypothesis, $n_0 \leqslant \binom{m-2}{l-2}$. Similarly, since
$$(a_{n_0},\ldots,\hat{a}_n) \in S(1,l,m-1), \qquad 2l \leqslant m-1,$$
we have $n-n_0 \leqslant \binom{m-2}{l-1}$. Thus
$$n = n_0+(n-n_0) \leqslant \binom{m-2}{l-2}+\binom{m-2}{l-1} = \binom{m-1}{l-1}.$$

Case 1 *b*. Suppose that there are $a \in A$, $\lambda \in [0,m)-a$ such that
$$m-1 \in a, \qquad a-\{m-1\}+\{\lambda\} \notin A.$$
Then $\lambda < m-1$. We may assume that
$$\begin{aligned}
&m-1 \in a_\nu, \quad \lambda \notin a_\nu, \quad b_\nu = a_\nu-\{m-1\}+\{\lambda\} \notin A \quad (\nu < n_0),\\
&m-1 \in a_\nu, \quad \lambda \notin a_\nu, \quad c_\nu = a_\nu-\{m-1\}+\{\lambda\} \in A \quad (n_0 \leqslant \nu < n_1),\\
&m-1 \in a_\nu, \quad \lambda \in a_\nu \quad (n_1 \leqslant \nu < n_2),\\
&m-1 \notin a_\nu \qquad\qquad\quad (n_2 \leqslant \nu < n).
\end{aligned}$$
Here $1 \leqslant n_0 \leqslant n_1 \leqslant n_2 \leqslant n$. Put $b_\nu = a_\nu$ $(n_0 \leqslant \nu < n)$. We now show that
$$(b_0,\ldots,\hat{b}_n) \in S(1,l,m). \tag{4}$$
Let $\mu < \nu < n$. We have to prove that
$$b_\mu \neq b_\nu, \qquad b_\mu b_\nu \neq \emptyset. \tag{5}$$
If $\mu < \nu < n_0$ or $n_0 \leqslant \mu < \nu$, then (5) clearly holds. Now let $\mu < n_0 \leqslant \nu$. Then $b_\mu \notin A$, $b_\nu = a_\nu \in A$, and hence $b_\mu \neq b_\nu$. If $n_0 \leqslant \nu < n_1$, then $c_\nu \in A$, and there is $\sigma \in a_\mu c_\nu$. Then $\sigma \neq \lambda$, $\sigma \neq m-1$, and $\sigma \in b_\mu b_\nu$. If $n_1 \leqslant \nu < n_2$, then $\lambda \in b_\mu b_\nu$. If, finally, $n_2 \leqslant \nu < n$, then there is

$\rho \in a_\mu a_\nu$. Then

$$\rho \in a_\nu, \qquad \rho < m-1, \qquad \rho \in b_\mu b_\nu.$$

This proves (5) and therefore (4). However, we have

$$f(b_0,..., \hat{b}_n)-f(a_0,..., \hat{a}_n) = n_0(-[m-1]+\lambda) < 0,$$

which contradicts the minimum property of $(a_0,..., \hat{a}_n)$. This shows that Case 1 b cannot occur.

Case 2: $\min(\nu < n)\ |a_\nu| = l_0 \leqslant l$.

If $l_0 = l$, then we have Case 1. Now let $l_0 < l$ and use induction over $l-l_0$. We may assume that

$$|a_\nu| = l_0 \quad (\nu < n_0), \qquad |a_\nu| > l_0 \quad (n_0 \leqslant \nu < n), \tag{6}$$

where $1 \leqslant n_0 \leqslant n$. Let $b_0,..., \hat{b}_{n_1}$ be the distinct sets b such that (3) holds for some ν. Then

$$(l_0+1)-(m-l_0) \leqslant 2(l-1)+1-m \leqslant 0, \tag{7}$$

and hence, by the lemma, $n_1 \geqslant n_0$. Also,

$$(b_0,..., \hat{b}_{n_1}, a_{n_0},..., \hat{a}_n) \in S(1, l, m).$$

Hence, from our induction hypothesis,

$$n \leqslant n_1+(n-n_0) \leqslant \binom{m-1}{l-1},$$

and Theorem 1 follows.

6. Proof of Theorem 2

Case 1: $k = 0$. Then

$$2l \leqslant 1+m, \qquad |a_\nu| \leqslant l \quad (\nu < n), \qquad (a_0,..., \hat{a}_n) \in S(0, l, m).$$

Now (a) (ii) is impossible, and (a) (i) is identical with (b), so that all we have to prove is that $n \leqslant \binom{m}{l}$. Again, we may assume (6), where $l_0 \leqslant l$, $1 \leqslant n_0 \leqslant n$. If $l_0 = l$ then

$$|a_\nu| = l \quad (\nu < n), \qquad n \leqslant \binom{m}{l}.$$

Now let $l_0 < l$ and use induction over $l-l_0$. Let $b_0,..., \hat{b}_{n_1}$ be the distinct sets b such that (3) holds for some ν. Again (7) holds and, by the lemma, $n_1 \geqslant n_0$. We have

$$(b_0,..., \hat{b}_{n_1}, a_{n_0},..., \hat{a}_n) \in S(0, l, m)$$

and, by our induction hypothesis,

$$n \leqslant n_1+(n-n_0) \leqslant \binom{m}{l}.$$

This proves the assertion.

Case 2: $k > 0$. We separate this into two cases.

Case 2*a*. Suppose that (1) holds. Put $|a_0 \ldots \hat{a}_n| = r$. We now show that, if $r \geqslant k$, then (i) follows. We may assume that

$$a_0 \ldots \hat{a}_n = [m-r, m).$$

Put
$$a_\nu[0, m-r) = c_\nu \quad (\nu < n).$$
Then
$$(c_0, \ldots, \hat{c}_n) \in S(0, l-r, m-r),$$
$$2(l-r)-(m-r) = 2l-r-m \leqslant 2l-k-m \leqslant 0.$$

Hence, by Case 1,

$$n \leqslant \binom{m-r}{l-r} = \binom{m-r}{m-l} \leqslant \binom{m-k}{m-l} = \binom{m-k}{l-k},$$

so that (i) holds. We now suppose that (i) is false, and we deduce (ii). We have
$$|a_0 \ldots \hat{a}_n| = r < k \leqslant |a_0 a_1| < |a_0| \leqslant l,$$
and therefore
$$2l \leqslant k+m < l+m, \qquad k < l < m.$$
There is a maximal number $p \geqslant n$ such that there exist $p-n$ sets $a_n, \ldots, \hat{a}_p$ satisfying
$$(a_0, \ldots, \hat{a}_p) \in S(k, l, m). \tag{8}$$
Put $A' = \{a_\nu : \nu < p\}$. We assert that
$$(a_0, \ldots, \hat{a}_p) \notin S(k+1, l, m). \tag{9}$$
For otherwise $|a_\mu a_\nu| > k$ $(\mu < \nu < n)$. Let $a \in A'$. Then we can choose $a' \subset [0, m)$ such that
$$|a'| = l, \qquad |aa'| = l-1.$$
Then, for every $b \in A'$, we have
$$|a'b| \geqslant |ab|-1 \geqslant k$$
and hence, since p is maximal, $a' \in A'$. By repeated application of this result we find that

$$[0, l), [m-l, m) \in A', \qquad k < |[0, l)[m-l, m)| = l-(m-l) \leqslant k,$$

which is the desired contradiction. This proves (9), and hence there are sets $a, b \in A'$ such that $|ab| = k$. Since $|a_0 \ldots \hat{a}_p| \leqslant |a_0 \ldots \hat{a}_n| < k$, there is $c \in A'$ such that $|abc| < k$. Denote by T the set of all triples (x, y, z) such that $x \subset a$, $y \subset b$, $z \subset c$, $|x| = |y| = |z| = k$, $|x+y+z| \leqslant l$. Put $\phi(x, y, z) = \{d : x+y+z \subset d \in A'\}$. Then, by (8),

$$A' = \sum ((x, y, z) \in T)\, \phi(x, y, z).$$

If $(x, y, z) \in T$ and $s = |x+y+z|$, then $s > k$ since otherwise we obtain the contradiction
$$k > |abc| \geqslant |xyz| = |x| = k.$$

Hence

$$|\phi(x,y,z)| \leqslant \binom{m-s}{l-s} = \binom{m-s}{m-l} \leqslant \binom{m-k-1}{m-l} = \binom{m-k-1}{l-k-1},$$

$$n \leqslant p = |A'| \leqslant \binom{m-k-1}{l-k-1}\binom{l}{k}^3,$$

which proves (ii).

Case 2 *b*. Suppose that (2) holds. We may assume (6), where $l_0 \leqslant l$; $1 \leqslant n_0 \leqslant n$. If $l_0 = l$, then Case 2 *a* applies. Now let $l_0 < l$ and use induction over $l-l_0$. Let $b_0,\ldots, \acute{b}_{n_1}$ be the distinct sets b satisfying, for some ν, the relations (3). Then (7) holds and hence, by the lemma, $n_1 \geqslant n_0$. Also, since $l_0 < l < m$, so that $m-l_0 \geqslant 2$, we have, by definition of the b_μ,

$$b_0 \ldots \acute{b}_{n_1} = a_0 \ldots \acute{a}_{n_0}, \qquad |b_0 \ldots \acute{b}_{n_1} a_{n_0} \ldots \acute{a}_n| = |a_0 \ldots \acute{a}_n| < k.$$

Since $$(b_0,\ldots, \acute{b}_{n_1}, a_{n_0},\ldots, \acute{a}_n) \in S(k,l,m),$$

it follows from our induction hypothesis that

$$n \leqslant n_1+(n-n_0) \leqslant \binom{m-k-1}{l-k-1}\binom{l}{k}^3.$$

It remains to prove (b) in Case 2. If $k = l$, then (b) is trivial. If $k < l$ and $m \geqslant k+(l-k)\binom{l}{k}^3$, then

$$\binom{m-k}{l-k} = \binom{m-k-1}{l-k-1}\frac{m-k}{l-k} \geqslant \binom{m-k-1}{l-k-1}\binom{l}{k}^3,$$

so that (b) follows from (a). This completes the proof of Theorem 2.

7. Concluding remarks

(i) In Theorem 2 (b) the condition

$$m \geqslant k+(l-k)\binom{l}{k}^3,$$

though certainly not best-possible, cannot be omitted. It is possible for

$$(a_0,\ldots, \acute{a}_n) \in S(k,l,m), \qquad k \leqslant l \leqslant m$$

to hold and, at the same time, $n > \binom{m-k}{l-k}$. This is shown by the following example due to S. H. Min and kindly communicated to the authors. Let $a_0,\ldots, \acute{a}_n$ be the distinct sets a such that

$$a \subset [0,8), \qquad |a| = 4, \qquad |a[0,4)| = 3.$$

Then

$$n = 16, \qquad (a_0,\ldots, a_{15}) \in S(2,4,8), \qquad \binom{m-k}{l-k} = \binom{6}{2} = 15 < n.$$

A more general example is the following. Let $r > 0$ and denote by $a_0, \ldots, \hat{a}_n$ the distinct sets a such that

$$a \subset [0, 4r), \qquad |a| = 2r, \qquad |a[0, 2r)| > r.$$

Then
$$(a_0, \ldots, \hat{a}_n) \in S(2, 2r, 4r),$$
and we have

$$n = \sum (r < \lambda \leqslant 2r) \binom{2r}{\lambda}\binom{2r}{2r-\lambda} = \tfrac{1}{2} \sum (\lambda \leqslant 2r) \binom{2r}{\lambda}\binom{2r}{2r-\lambda} - \frac{1}{2}\binom{2r}{r}^2$$
$$= \frac{1}{2}\binom{4r}{2r} - \frac{1}{2}\binom{2r}{r}^2.$$

In this case
$$\binom{m-k}{l-k} = \binom{4r-2}{2r-2},$$
and, for every large r, possibly for every $r > 2$, we have $\binom{m-k}{l-k} < n$.

We put forward the conjecture that, for our special values of k, l, m, this represents a case with maximal n, i.e.

If $\qquad r > 0, \qquad (a_0, \ldots, \hat{a}_n) \in S(2, 2r, 4r),$

then
$$n \leqslant \frac{1}{2}\binom{4r}{2r} - \frac{1}{2}\binom{2r}{r}^2.$$

(ii) If in the definition of $S(1, l, m)$ in § 2, the condition $a_\mu \not\subset a_\nu \not\subset a_\mu$ is replaced by $a_\mu \neq a_\nu$ and if no restriction is placed upon $|a_\nu|$, then the problem of estimating n becomes trivial, and we have the result:

Let $m > 0$ *and* $a_\nu \subset [0, m)$ *for* $\nu < n$, *and* $a_\mu \neq a_\nu$, $a_\mu a_\nu \neq \emptyset$ *for* $\mu < \nu < n$. *Then* $n \leqslant 2^{m-1}$, *and there are* $2^{m-1} - n$ *subsets* $a_n, \ldots, \hat{a}_{2^{m-1}}$ *of* $[0, m)$ *such that* $a_\mu \neq a_\nu$, $a_\mu a_\nu \neq \emptyset$ *for* $\mu < \nu < 2^{m-1}$.

To prove this we note that of two sets which are complementary in $[0, m)$ at most one occurs among $a_0, \ldots, \hat{a}_n$, and, if $n < 2^{m-1}$, then there is a pair of complementary sets a, b neither of which occurs among $a_0, \ldots, \hat{a}_n$. It follows that at least one of a, b intersects every a_ν, so that this set can be taken as a_n.

(iii) Let $l \geqslant 3$, $2l \leqslant m$, and suppose that

$$a_\nu \subset [0, m), \qquad |a_\nu| = l \quad \text{for} \quad \nu < n,$$
and
$$a_\mu \neq a_\nu, \qquad a_\mu a_\nu \neq \emptyset \quad \text{for} \quad \mu < \nu < n, \quad \text{and} \quad a_0 \ldots \hat{a}_n = \emptyset.$$

We conjecture that the maximum value of n for which such sets a_ν can be found is n_0, where

$$n_0 = 3\binom{m-3}{l-2} + \binom{m-3}{l-3}.$$

A system of n_0 sets with the required properties is obtained by taking all sets a such that

$$a \subset [0,m), \qquad |a[0,3)| \geqslant 2, \qquad |a| = l.$$

(iv) The following problem may be of interest. Let $k \leqslant m$. Determine the largest number n such that there is a system of n sets a_ν satisfying the conditions

$$a_\mu \neq a_\nu, \qquad |a_\mu a_\nu| \geqslant k \quad (\mu < \nu < n).$$

If $m+k$ is even, then the system consisting of the a such that

$$a \subset [0,m), \qquad |a| \geqslant \tfrac{1}{2}(m+k)$$

has the required properties. We suspect that this system contains the maximum possible number of sets for fixed m and k such that $m+k$ is even.

(v) If in (ii) the condition $a_\mu a_\nu \neq \emptyset$ $(\mu < \nu < n)$ is replaced by $a_\mu a_\nu a_\rho \neq \emptyset$ $(\mu < \nu < \rho < n)$, then the structure of the system a_ν is largely determined by the result:

Let $m \geqslant 2$, $a_\nu \subset [0,m)$ *for* $\nu < n$, $a_\mu \neq a_\nu$ *for* $\mu < \nu < n$, *and* $a_\mu a_\nu a_\rho \neq \emptyset$ *for* $\mu < \nu < \rho < n$. *Then* $n \leqslant 2^{m-1}$, *and, if* $n = 2^{m-1}$, *then* $a_0 a_1 \ldots \hat{a}_n \neq \emptyset$, *so that the* a_ν *are all* 2^{m-1} *sets* $a \subset [0,m)$ *which contain some fixed number* t $(t < m)$.

For there is a largest p $(1 \leqslant p \leqslant n)$ such that

$$a_0 a_1 \ldots \hat{a}_p \in \{a_0, a_1, \ldots, \hat{a}_n\}.$$

If $p = n$, then $a_0 \ldots \hat{a}_n = a_\nu \neq \emptyset$ for some $\nu < n$. If $p < n$, then any two of the $n+1$ distinct sets

$$a_0 a_1 \ldots a_p,\ a_0,\ a_1, \ldots,\ \hat{a}_n$$

have a non-empty intersection and hence, by (ii), $n+1 \leqslant 2^{m-1}$. Different proofs of (v) have been found by L. Pósa, G. Hajós, G. Pollák, and M. Simonovits.

REFERENCE

1. E. Sperner, *Math. Z.* 27 (1928) 544–8.